Full Duplex Communication in Cognitive Radio Networks

Towhidlou, Vahid

Awarding institution:
King's College London

FULL DUPLEX COMMUNICATION IN COGNITIVE RADIO NETWORKS

By Vahid Towhidlou

A thesis submitted to King's College London

In fulfilment of the requirements for the Degree of

Doctor of Philosophy

May 2016

ABSTRACT

Over the past decade mobile services have evolved from basic voice communication to mobile-broadband multimedia services and sophisticated wireless applications. Today, we cannot imagine a world without wireless communications and smart mobile devices. Any wireless communication relies on the availability of radiofrequency spectrum, which is inherently a finite resource and cannot be produced. This unprecedented growth of wireless devices and services has motivated academia and industry to look for new solutions in utilizing the wireless spectrum more intelligently and efficiently.

Cognitive radio (CR) and Full Duplex (FD) communication are two promising technologies recently developed to enhance spectrum utilization and network efficiency, and combination thereof will improve the performance even further. CR technology has been extensively studied within the past 15 years and made available for commercial applications. On the other hand, full duplex communication was not deemed possible due to some complications until recently, when the advances in Self-Interference Suppression (SIS) methods made it possible to implement this technology in future wireless networks.

In line with the global attention on full duplex and cognitive radio technologies, this research investigates new schemes of implementing full duplex communication in cognitive radio networks. In part of this thesis, the advantages of in-band full duplex transmissions in doubling the spectral efficiency and enhancing overall performance of a system, have been studied in a proposed scheme incorporating cooperative Automatic Repeat reQuest (ARQ) in an overlay cognitive radio network. It is shown that a full duplex cooperative ARQ protocol will increase the throughput of primary and secondary networks, without any degradation on primary network performance.

Another part of this thesis focuses on asynchronous bidirectional FD communication in a CR system. While detection probability is enhanced due to

cooperative sensing, the probability of collision and average duration of collisions are decreased significantly thanks to the implementation of asynchronous full duplex transmissions. These gains in network performance are additional to improved throughput of secondary network achieved through implementation of full duplex transmission. The proposed schemes are mathematically analyzed and have been numerically verified through extensive simulations.

As a contribution to IEEE 802.22 Standard revision, a novel suggestion incorporating the use of directional sector antennas, and full duplex operation of secondary users in a Wireless Regional Area Network (WRAN) has been proposed in another part of this thesis and it has been shown how research achievements may be implemented in practice.

A summary of the thesis plus some suggestions for future work conclude this thesis.

Table of Contents

Table of Figures

Table of Tables

Acknowledgements

First of all, I would like to express my genuine gratitude and appreciation to my supervisor, Dr. Mohammad Shikh-Bahaei for his support, encouragement, and contributions in the development of my research. This thesis would not have been possible without his advice, attention and cooperation.

I would like to take this opportunity to thank my dear friends Mr. Sobhan Yassipour Tehrani and Mr. Seyed Ali Hosseini for their sincere support and friendship, and for the happy times we spent together throughout our journey at King's. I also acknowledge all my lab colleagues and friends for making the work environment at King's friendly and lovely.

Last, but not least, I would like to express my special appreciation and thanks to my dear parents for their love, sacrifice and support. This thesis is dedicated to them.

Abbreviations

5G	Fifth Generation
ACK	Acknowledge
ADC	Analog to Digital Convertor
AF	Amplify and Forward
AMC	Adaptive Modulation and Coding
AP	Access Point
APP	Application
ARQ	Automatic Repeat Query; Automatic Retransmission reQuest
AWGN	Additive White Guassian Noise
BER	Bit Error Rate
BF	Beam Forming
BKIC	Blind Known Interference Cancellation
BLP	Basic Linear Programming
BPSK	Bipolar Phase Shift Keying
CA	Carrier Aggregation
CAF	Cyclic Autocorrelation Function
CBS	Cognitive Base Station
CC	Component Carrier
CL	Cross Layer
CLT	Central Limit Theorem
CPE	Customer Premises Equipment
CR	Cognitive Radio

CRN	Cognitive Radio Network
CS	Cooperative Sensing
CSD	Cyclic Spectral Density
CSI	Channel State Information
CTS	Clear To Send
CW	Continuous Wave
dB	Decibel
DF	Decode and Forward
DLL	Data Link Layer
DSA	Dynamic Spectrum Access
FBMC	Filter Bank Multi Carrier
FCC	Federal Communication Commission
FD	Full Duplex
FDCR	Full Duplex Cognitive Radio
FDD	Frequency Division Duplex
FDTR	Full Duplex Transmit and Receive
FER	Frame Error Rate
FIFO	First In First Out
HARQ	Hybrid ARQ
HD	Half Duplex
HIPERLAN	High Performance Local Area Network
HSDPA	High Speed Downlink Packet Access
HSPUA	High Speed Uplink Packet Access
IBFD	In-Band Full Duplex

IC	Interference Cancellation
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
ISM	Industrial, Scientific, and Medical
KIC	Known Interference Cancellation
LAN	Local Area Network
LBT	Listen Before Talk
LNA	Low Noise Amplifier
LP	Linear Programming
LTE	Long Term Evolution
LTE-A	LTE Advanced
M2M	Machine to Machine
MAC	Medium Access Control
MHz	Mega Hertz
MIMO	Multiple Input Multiple Output
MINLP	Mixed Integer Non-Linear Programming
MMSE	Minimum Mean Square Error
NACK	Not Acknowledge
NOMA	Non Orthogonal Multiple Access
NSP	Null Space Projection
Ofcom	Office of Communications
OSA	Opportunistic Spectrum Access
pdf	probability density function
PER	Packet Error Rate

PHY	Physical Layer
PSD	Power Spectral Density
PSR	Packet Success Rate
PU	Primary User
QoS	Quality of Service
RF	Radio Frequency
RIB	Rate Interference Budget
RTS	Ready To Send
SA	Spectrum Allocation/Assignment
SDR	Software Defined Radio
SHARP	Spectrum Harvesting with ARQ Retransmission and Probing
SIC	Self-Interference Cancellation
SINR	Signal to Interference plus Noise Ratio
SIS	Self-Interference Suppression
SNR	Signal-to-Noise Ratio
STBC	Space Time Block Code
SU	Secondary User
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDM	Time Division Multiplexing
TR	Transmission and Reception
TVWS	TV White Space
UMTS	Universal Mobile Telecommunications Service
UWB	Ultra Wide Band

VoIP	Voice over Internet Protocol
WiMax	Worldwide Interoperability for Microwave Access
WRAN	Wireless Regional Area Network
WSS	Wide Sense Stationary
ZF	Zero Forcing

Chapter 1

Introduction

1.1 Background and Motivation

It would be hard to imagine a world without wireless devices, applications, and services. Around the globe, mobile services are playing increasingly important roles in many facets of our society. During the past 10 years, mobile services have evolved from basic voice communication to mobile-broadband multimedia services. Today, we depend on mobile services not only for communication, but also for education, entertainment, healthcare, location and m-commerce [1]. Advancements in connected cars, smart grids, machine-to-machine (M2M) communication, and domestic installations such as at-home health monitoring systems, are some examples of emerging wireless services of tomorrow. Mobile data traffic has grown 4,000-fold over the past 10 years and almost 400-million-fold over the past 15 years and global mobile data traffic will increase nearly eightfold between 2015 and 2020 [2].

Any wireless communication relies on the availability of radiofrequency spectrum, which is inherently a finite resource and cannot be produced. In many global cities, we are rapidly approaching the point of "peak data", where user demand for wireless services can no longer be fully accommodated by the available radiofrequency spectrum. This unprecedented growth in wireless devices and mobile data traffic, and corresponding demand for wireless bandwidth, has motivated academia and industry to look for new solutions in utilizing the wireless spectrum more intelligently and efficiently. Although there have been numerous efforts and studies with some outstanding results over the past decade, the quest for new communication and networking paradigms to manage and utilize the limited radio spectrum as efficient as possible, is however felt more than ever.

Generally, there have been two main approaches in studies focusing on efficient utilization of spectrum. The first approach involves sharing of the spectrum of licensed or unlicensed networks by secondary users, and the other approach involves systems based on utilizing advanced air interface transmission technologies, such as advanced interference mitigation techniques, massive multiple-input multiple-output (MIMO) networks, and simultaneous non-orthogonal transmission methods.

Among the first approach technologies, the notion of Cognitive Radio (CR) is the most important paradigm which has attracted much attention within the past decade and has been recently standardized for commercial applications [3]. The concept of cognitive radio is based on an intelligent system which interacts with its environment and dynamically and autonomously adjust its operational parameters and protocols according to its obtained knowledge in order to achieve predefined objectives [4]. One of these main objectives is the shared exploitation of a spectrum band by licensed (primary) and unlicensed (secondary) users which is referred to as the notion of spectrum sharing. Within the operation of a primary network, there are some unused portions of licensed spectrum in time or frequency domain, which are potential opportunities to be used by secondary users. As the name implies, primary users have priority in using the bandwidth, hence secondary users should be intelligent enough to detect such opportunities and dynamically access them without impinging any interference on the performance of the primary networks.

Among second approach methods, (i.e. non-orthogonal transmission), In-Band Full-Duplex (IBFD) transmission (simultaneous transmission and reception over the same frequency band), Non-Orthogonal Multiple Access (NOMA) [5], [6], and non-orthogonal Filter Bank Multi Carrier (FBMC) [6], have recently attracted significant attention in academia and industry because they have the potential of being able to increase spectral efficiency without the need for additional frequency resources. Of these, IBFD transmission has a long history and has been used in the design of continuous wave (CW) radar systems since the 1940s [7]. But until recently it has not come into widespread use due to the severe effects of its inherent self-interference. Recent advances in self-interference cancellation techniques have shown the feasibility of IFDB communications, and has attracted significant interest in re-architecting wireless communication systems based on IFDB technology [6].

Apart from these two main approaches, there are some methods for aggregation and defragmentation of unused or unallocated frequency bands for the use of high data-rate applications. In this approach which is called Carrier Aggregation (CA) in the literature, two or more blocks of spectrum, or Component Carriers (CCs) are combined together to create wider channel bandwidths, which helps achieve much-needed faster data rates for new generation mobile technologies such as LTE Advanced (LTE-A) [8]. Through this method smaller carriers which may not be used for high speed applications in conventional ways are combined together and an efficient exploitation of spectrum is achieved.

Since the above paradigms are not in contradiction or overlapping with each other, they may be combined together to increase the spectral efficiency even more. For example, a cognitive radio network may detect more than one unused (free) spectrum slot at a time which may be aggregated to create a wider bandwidth for the use of a secondary network (CR and CA combined). Or the cognitive capability of a secondary wireless network may be employed to identify any potential transmission opportunity, and then full duplex transmissions to achieve double spectral efficiency (CR and FD combined). In addition to boosting spectral efficiency, joint paradigms may improve the overall performance of a wireless system in terms of lower error rates, enhanced QoS, and security and secrecy.

Recent advances in FD communication practicality in wireless systems, and the perspective of the significant role of CR technology in new generation mobile systems and services have recently motivated many researchers to focus on Full Duplex Cognitive Radio (FDCR) systems as the main wireless technology of the future [9], [10]. In this research, in line with the research community, we have studied the principles of FD and CR paradigms, and focused on a few practical schemes incorporating FDCR technology.

1.2 Thesis Contributions

In this thesis we have considered implementation of FD communications in a CR network and proposed two novel schemes in which not only the spectral efficiency of the overall network is enhanced, but also the other advantages of application of FD technology improve crucial metrics of the whole system as well. It is shown that implementation of full duplex transmissions within a cooperative automatic repeat request (ARQ) protocol in a cognitive radio system will not only increase the throughput of both primary and secondary networks, but will also decrease the error rate in the primary network at the same time; an achievement not possible in half duplex systems. For this purpose, we have proposed a novel cooperative protocol based on full duplex capability of secondary nodes which takes advantage of the opportunities arising during primary ARQ retransmission intervals. We have analyzed the packet error rate and throughput of the primary and secondary networks mathematically for ARQ and Hybrid ARQ schemes and have shown that this protocol not only does not harm the throughput of the primary, but will also enhance it in a Hybrid ARQ scheme due to incorporating a cooperative scheme. In contrast to other cooperative methods based on network coding, this method is easy to implement without any complexity, and higher throughputs are achievable for both primary and secondary systems. In addition, in this scheme PUs are oblivious to SUs and will benefit from the secondary's cooperation whenever applicable.

In another chapter, it is shown how asynchronous full duplex communication in a cognitive radio network may double spectral efficiency of the cognitive platform, while reducing the probability and duration of collisions with the primary signal. We have presented a novel scheme for full duplex communication in CR networks which consists of cooperative sensing and asynchronous FD transmissions by secondary users. We have analyzed the impact of these methods on network metrics such as probability of false alarm and detection, collision probability and duration, and the CR network average throughput. We have derived closed forms for the probability density function of the collision duration in a CR network where the activity of PUs is modeled by an ON/OFF Markov process, and channels are Rayleigh flat fading. In addition, we have derived analytical closed forms for average throughput of the secondary network under such conditions and have shown that our scheme outperforms the traditional methods in terms of cognitive network metrics.

In order to show how cognitive radio and full duplex paradigms may be combined together in practice, we have presented a new idea as a contribution to IEEE 802.22 WRAN standard [3]. This idea suggests the exploitation of directional sector antennas with or without full duplex capability to improve secondary network's performance in a regional area network. It has been shown that application of directional sector antennas at cognitive network's Base Station (BS) will equip the secondary network with directional detection of the primary signal presence, and may result in higher secondary network's throughput. When the cognitive BS has the capability to operate in full duplex mode, it can transmit and receive in opposite directions to different Customer Premises Equipments (CPEs) at the same time, thus increasing the network throughput, as well as fairness between CPEs.

In the appendix, a new model for Rayleigh flat fading channel based on queuing theory is presented and analyzed. It is shown that any wireless transmission system may be modeled by a combination of series and parallel exponential queues which provides us with a very useful tool for predicting the system's behavior, and meeting Quality of Service (QoS) requirements in a wireless system. As an instance, the Rayleigh channel has been considered and modeled by a hyperexponential queue and its characteristics (such as waiting time and service time distributions) have been analyzed and verified by numerical simulations. Such tool may be beneficial in analysis of any wireless system such as cognitive radio, especially when QoS is of great importance.

1.3 Thesis Structure

The remainder of this thesis is organized as follows:

Chapter 2 is dedicated to providing an overview of cognitive radio and full duplex communication paradigms as two promising technologies for better utilization of radio spectrum, as well as FDCR for next generation mobile systems. Respective literature review accompanies each section of this chapter.

Chapters 3 and 4 are the main body of this thesis, and in each chapter a novel scheme of full duplex communication in a cognitive radio network is proposed and studied. Chapter 3 proposes a cooperative ARQ scheme for an overlay cognitive radio network with full duplex capability. Performance of the system model for ARQ and Hybrid ARQ schemes is investigated and verified by extensive simulations.

In Chapter 4 an interweave cognitive radio network has been considered in which secondary users are capable of partial self-interference suppression and operate in bidirectional full duplex mode. Cooperative sensing has been proposed for improving detection probability and providing better protection for non-time-slotted primary network operation.

Chapter 5 proposes a novel idea submitted to IEEE 802.22 Standard committee as a contribution to standard revision. The proposed idea shows how implementation of directional sector antenna and FD capability in the base station of a WRAN will improve the whole network's performance in practice.

Finally, Chapter 6 provides concluding remarks on the entire research conducted and highlights some suggestions for future works and improvements.

1.4 Research Contributions

The contributions and novelties of this thesis have been drawn from and are disseminated through the following technical papers:

[C1] V. Towhidlou, M. Shikh-Bahaei,” Rayleigh flat fading channel modeled as $G/H_R/1$ queue”, In Proc. IEEE European Conference on Networks and Communications, EuCNC 2015, pp. 41-45.

[C2] V. Towhidlou, M. Shikh-Bahaei, “ Asynchronous full duplex cognitive radio”, to appear in the proceedings of IEEE 84th Vehicular Technology Conference (VTC 2016-Fall).

[C3] V. Towhidlou, M. Shikh-Bahaei,” Cooperative ARQ in full duplex cognitive radio networks” in Proc. IEEE 27th annual international symposium on personal, indoor, and mobile radio communications (PIMRC 2016).

[C4] V. Towhidlou, M. Shikh-Bahaei,” Full Duplex Cooperative ARQ and HARQ in cognitive radio networks” submitted to IEEE ICC Conference 2017.

[J1] V. Towhidlou, M. Shikh-Bahaei,” Cross layer design in cognitive radio networks: a survey” submitted to IEEE communication magazine, to appear.

[J2] V. Towhidlou, M. Shikh-Bahaei, “Full duplex cooperative ARQ and HARQ in cognitive radio networks” to appear in IEEE Wireless Communication Letters.

[J3] V. Towhidlou, M. Shikh-Bahaei, “Asynchronous Full-Duplex Cognitive Networks: Physical and MAC layer analysis” submitted to IEEE Transactions on Cognitive Communications and Networking.

[SC] V. Towhidlou, M. Shikh-Bahaei, O. Holland, M. Dohler, and H. Aghvami,”Directional antennas; full duplex communication” contribution submitted to IEEE 802.22 standard committee, doc.: IEEE 802.22-16/0033r0, Nov. 2016.

Chapter 2

Full Duplex Cognitive Radio

2.1 Cognitive Radio

Studies of the spectrum scarcity problem by various regulatory bodies around the globe, including the Federal Communication Commission (FCC) in the United States of America and OfCom in the United Kingdom have indicated that the licensed spectrum bands are severely under-utilized at any given time and location [11] and the overall spectrum utilization is less than 15%, mainly due to the traditional static spectrum allocation regulation [11]. Under such a spectrum allocation, the radio spectrum is divided into fixed bands for exclusive use of license holder operators for a specific type of service and radio device. Hence, to address the spectrum underutilization and spectrum scarcity problems, regulatory bodies decided to allow unlicensed users to opportunistically access licensed bands when these bands are not occupied by their primary holders. This decision stemmed from a suggestion called "Cognitive Radio".

The concept of CR, first proposed in 1999 by Joseph Mitola III [12], opened up a new way of better utilization of the scarce wireless spectrum resource. In fact, the original idea introduced by Mitola embraced a broader perspective, and efficient utilization of wireless spectrum is one of the applications of a cognitive system;

"*The point in which wireless personal digital assistants (PDAs) and the related networks are sufficiently computationally intelligent about radio resources and related computer-to-computer communications to detect user communications needs as a function of use context, and to provide radio resources and wireless services most appropriate to those needs*" [12].

In other words, CR is defined as an intelligent radio that can be programmed and configured dynamically based on interaction with its environment to achieve predefined objectives, one of which can be shared use of a spectrum band. As an intelligent radio, it has the *Cognitive Capability* to identify any transmission opportunity in time, frequency, and space domains through spectrum sensing, and accordingly reconfigure its transceiver parameters to communicate over the identified opportunities (*Reconfigurability*) in a way to inflict minimum interference on the operation of other networks sharing the same spectrum.

Dynamic Spectrum Access (DSA) is such an important and well-studied application of cognitive radio technology that the terms "cognitive radio" and "DSA" are almost used interchangeably in wireless communication literature. In this thesis as well, we are focusing on DSA applications of cognitive radio technology and wherever we speak of CR, we mean the spectrum sharing and DSA application of a cognitive radio system.

In DSA systems, unlicensed or secondary users (SUs) were allowed to identify the under-utilized portions of licensed spectrum, often called "white holes" or "spectrum holes", and utilize them opportunistically as long as they do not cause any harmful interference to the communications of the licensed or primary users (PUs). Later on, this paradigm was extended to identification of any opportunity available within primary network operation and taking advantage of it as long as the interference impinged on the primary network is within tolerable limits. Since its introduction, cognitive radio has been regarded as a promising technology to increase spectrum utilization through spectrum sharing between licensed and unlicensed users and has attracted much attention from industry and academia. A few years ago it was standardized for commercial applications and implemented for Wireless Regional Area Networks (WRANs) to provide internet access for remote areas [3].

Based on the type of available network side information along with the regulatory constraints, secondary users seek to underlay, overlay, or interweave their signals with those of primary users without significantly impacting these users. In the next section we describe these different cognitive radio paradigms in more detail.

2.1.1 Cognitive Radio Network Paradigms

There are three main cognitive radio network paradigms: underlay, overlay, and interweave[1]. In the underlay paradigm, primary and secondary users coexist and operate simultaneously. This paradigm allows secondary users to operate if the interference they cause to primary users is below a given threshold or meets a given bound on primary user performance degradation. Rather than determining the exact interference it causes, a secondary user can spread its signal over a very wide bandwidth such that the interference power spectral density is below the noise floor at any primary user location [13]. These spread signals are then despread at each of their intended secondary receivers. This spreading technique is the basis of both spread spectrum and ultra wide band (UWB) communications. Alternatively, the secondary transmitter can be very conservative in its output power to ensure that its signal remains below the prescribed interference threshold. In this case, since the interference constraints in underlay systems are typically quite restrictive, this limits the secondary users to short range communications [13].

In overlay systems the secondary users overhear the transmissions of the primary users, then use this information along with sophisticated signal processing and coding techniques to maintain or improve the performance of primary users, while also

[1] Sometimes underlay and interweave, or overlay and interweave paradigms are applied together in a CR scheme, which are referred to "mixed paradigm" in some texts.

obtaining some additional bandwidth for their own communication. The premise for overlay systems is that the secondary transmitter has knowledge of the primary user's transmitted data sequence (also called its message) and how this sequence is encoded (also called its codebook). Similar ideas apply when there are multiple secondary and primary users. The codebook information could be obtained, for example, if the primary users follow a uniform standard for communication based on a publicized codebook. Alternatively, the primary users could broadcast their codebooks periodically. A primary user's data sequence might be obtained by decoding it at the secondary user's receiver or in other ways. Knowledge of a primary user's data sequence and/or codebook can be exploited in a variety of ways to either cancel or mitigate the interference seen at the secondary and primary receivers. On the one hand, this information can be used to cancel the interference due to the primary signals at the secondary receiver. Specifically, sophisticated encoding techniques can be used to precode the secondary user's signal such that the known primary user interference at the secondary receiver is effectively removed. On the other hand, the secondary users can assign part of their power for their own communication and the remainder of the power to assist (relay) the primary transmissions. By careful choice of the power split, the increase in the primary user's signal-to-interference-plus-noise power ratio (SINR) due to the cooperation with secondary users can be exactly offset by the decrease in the primary user's SINR due to the interference caused by the fraction of the secondary user's power assigned to its own communication. Under ideal conditions, sophisticated encoding and decoding strategies allow both the secondary and primary users to remove all or part of the interference caused by other users [13].

In interweave systems the secondary users detect the absence of primary user signals in space, time, or frequency, often named as white holes or spectrum holes, and opportunistically communicate during these absences. The interweave technique requires detection of primary (licensed or unlicensed) users in one or more of the space-time frequency dimensions. This detection is quite challenging since primary user activity changes over time and also depends on geographical location. If detection is not perfect, as is the case in real scenarios, the secondary user's signal will collide with the present or reappeared primary's signal and this interference will degrade the performance of both networks.

For all three paradigms, if there are multiple secondary users then these users must share bandwidth amongst themselves as well as with the primary. This may be accomplished through a centralized sharing protocol, or a contention based sharing scheme in which secondary users contend for accessing the available transmission opportunity.

2.1.2 Cognitive Radio Functions

As explained in the previous section a CR manages the available detected transmission opportunities in licensed spectrum for its own use in a way not to interfere with primary users. The spectrum management process consists of four major functions:

2.1.2.1 Spectrum Sensing

Spectrum sensing is one of the main components of the CR concept. It can be conceptualized as the task of gaining awareness about the spectrum usage and existence of PUs in the surroundings of a CR-enabled transceiver. Signal detection is the core element of CR sensing and many sensing techniques have been proposed over the last decade based on properties of the signals to be detected. Here we explain some of the widely-used methods.

A) Energy Detection

Energy detection based approach, also known as radiometry or periodogram, is the most common way of spectrum sensing because of its low computational and implementation complexities [14], [15] [16]. In addition, it is more generic (as compared to methods given in this section) as receivers do not need any knowledge on the primary users' signal. The signal is detected by comparing the output of the energy detector with a threshold which depends on the noise floor. Some of the challenges with energy detector based sensing include selection of the threshold for detecting primary users, inability to differentiate interference from primary users and noise, and poor performance under low signal-to-noise ratio (SNR) values [17]. Moreover, energy detectors do not work efficiently for detecting spread spectrum signals [18], [19]. Let us assume that the received signal has the following simple form:

$$y(n) = s(n) + w(n), \tag{2.1}$$

where *s(n)* is the signal to be detected, *w(n)* is the additive white Gaussian noise (AWGN) sample, and *n* is the sample index. Note that *s(n)=0* when there is no transmission by the primary user. The decision metric for the energy detector can be written as:

$$M = \frac{1}{N}\sum_{n=0}^{N} |y(n)|^2, \tag{2.2}$$

where *N* is the number of samples. For simplicity we assume $N = f_s \cdot T_s$ in which f_s is the frequency of sampling and T_s is the duration of sampling slot. The decision on the occupancy of a band can be obtained by comparing the decision metric M against a fixed threshold ϵ. This is equivalent to distinguishing between the following two hypotheses:

$$\begin{aligned} \mathcal{H}_0 \quad &: \quad y(n) = w(n), \\ \mathcal{H}_1 \quad &: \quad y(n) = s(n) + w(n). \end{aligned} \tag{2.3}$$

The performance of the detection algorithm can be summarized with two probabilities; probability of detection P_d and probability of false alarm P_f. P_d is the probability of detecting a signal on the considered frequency when it truly is present. Thus, a large detection probability is desired. It can be formulated as:

$$P_d = \Pr(M > \epsilon | \mathcal{H}_1). \tag{2.4}$$

P_f is the probability that the test incorrectly decides that the considered frequency is occupied when it actually is not, and it can be written as:

$$P_f = \Pr(M > \epsilon | \mathcal{H}_0). \quad (2.5)$$

P_f should be kept as small as possible in order to prevent underutilization of transmission opportunities. The decision threshold ϵ can be selected for finding an optimum balance between P_d and P_f. However, this requires knowledge of noise and detected signal powers. The noise power can be estimated, but the signal power is difficult to estimate as it changes depending on ongoing transmission characteristics and the distance between the cognitive radio and primary user. In practice, the threshold is chosen to obtain a certain false alarm rate [20]. Hence, knowledge of noise variance is sufficient for selection of a threshold.

The white noise can be modeled as a zero-mean Gaussian random variable with variance σ_w^2, i.e. $w(n) = \mathcal{N}(0, \sigma_w^2)$. For a simplified analysis, let us model the signal term as a zero-mean Gaussian variable as well, i.e. $s(n) = \mathcal{N}(0, \sigma_s^2)$.The model for $s(n)$ is more complicated as fading should also be considered. Because of these assumptions, the decision metric (2.2) follows Chi-square distribution with 2N degrees of freedom χ_{2N}^2 and hence, it can be modeled as:

$$M = \begin{cases} \dfrac{\sigma_w^2}{2}\chi_{2N}^2 & \mathcal{H}_0, \\ \dfrac{\sigma_w^2 + \sigma_s^2}{2}\chi_{2N}^2 & \mathcal{H}_1. \end{cases} \quad (2.6)$$

For circularly symmetric complex Guassian (CSCG) noise and signal, the probabilities P_f and P_d can be approximated as:

$$P_f(\epsilon) = \mathcal{Q}\left(\left(\frac{\epsilon}{\sigma_w^2} - 1\right)\sqrt{N}\right), \quad (2.7)$$

$$P_d(\epsilon) = \mathcal{Q}\left(\left(\frac{\epsilon}{\sigma_w^2} - \gamma - 1\right)\sqrt{\frac{N}{2\gamma + 1}}\right), \quad (2.8)$$

where γ is the SNR of the signal to be detected and $\mathcal{Q}(\cdot)$ is the complementary distribution function of the standard Gaussian, or complementary error function, i.e.,

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty exp\left(-\frac{t^2}{2}\right)dt. \quad (2.9)$$

B) Cyclostationarity detection

Cyclostationarity detection is a method for detecting primary user transmissions by exploiting the cyclostationarity features of the received signals. Cyclostationary features are caused by the periodicity in the signal or in its statistics like mean and autocorrelation [21] or they can be intentionally induced to assist spectrum sensing [22], [23]. Instead of Power Spectral Density (PSD), cyclic correlation function is used for detecting signals present in a given spectrum. The cyclostationarity based detection algorithms can differentiate noise from primary users' signals. This is a result of the fact that noise is Wide-Sense Stationary (WSS) with no correlation while modulated signals are cyclostationary with spectral correlation due to the redundancy of signal periodicities [24] and cyclostationarity detection capitalizes on this property to detect a PU signal even at low SNRs. Furthermore, cyclostationarity can be used for distinguishing among different types of transmissions and primary users [25]. The Cyclic Spectral Density (CSD) function of a received signal (2.1) can be calculated as [21]:

$$S(f,\alpha) = \sum_{\tau=-\infty}^{\infty} R_y^{\alpha}(\tau)e^{-j2\pi f\tau}, \tag{2.10}$$

where

$$R_y^{\alpha}(\tau) = E[y(n+\tau)y^*(n-\tau)e^{j2\pi\alpha n}, \tag{2.11}$$

is the Cyclic Autocorrelation Function (CAF) and α is the cyclic frequency. The cyclic spectral density function outputs peak values when the cyclic frequency is equal to the fundamental frequencies of transmitted signal x(n). Cyclic frequencies can be assumed to be known [26] or they can be extracted and used as features for identifying transmitted signals [27].

C) Waveform-based detection

The waveform-based detection method is only applicable to systems with known signal patterns, and it is termed as waveform-based sensing or coherent sensing. Known patterns are usually utilized in wireless systems to assist synchronization or for other purposes. Such patterns include preambles, midambles, regularly transmitted pilot patterns, spreading sequences etc. A preamble is a known sequence transmitted before each burst and a midamble is transmitted in the middle of a burst or slot. In the presence of a known pattern, sensing can be performed by correlating the received signal with a known copy of itself [17], [28]. In [17] it is shown that waveform-based sensing outperforms energy detector based sensing in reliability and convergence time. Furthermore, it is shown that the performance of the sensing algorithm increases as the length of the known signal pattern increases.

D) Matched filtering

Matched filtering is known as the optimum method for detection of primary users when the transmitted signal is known [74]. The main advantage of matched filtering is the short time to achieve a certain probability of false alarm or probability of misdetection [29] as compared to other methods that are discussed in this section. In fact, the required number of samples grows as $O(1/SNR)$ for a target probability of false alarm at low SNRs for matched filtering [29]. However, matched-filtering requires cognitive radio to demodulate received signals. Hence, it requires perfect knowledge of the primary users signaling features such as bandwidth, operating frequency, modulation type and order, pulse shaping, and frame format. Moreover, since cognitive radio needs receivers for all signal types, the implementation complexity of sensing unit is impractically large. Another disadvantage of match filtering is large power consumption as various receiver algorithms need to be executed for detection.

E) Other Sensing Methods

Other alternative spectrum sensing methods include multitaper spectral estimation [30], wavelet transform based estimation [31], radio identification based sensing [19], Hough transform, and time-frequency analysis [32], which are not explained in this thesis and more information about them can be found in respective references.

In addition to signal detection which clarifies presence or absence of a primary user, spectrum sensing also involves acquiring some information about the primary's transmission parameters such as the modulation scheme, waveform, bandwidth and carrier frequency, codebook, which are essential to operation in overlay paradigms. Cooperative spectrum sensing has also been proposed as a means of improving the sensing accuracy and relaxing the sensing burdens of each individual CR.

2.1.2.2 Spectrum Allocation

After detection of available bands and some of their characteristics through the sensing cycle, now a CR should decide on the best spectrum band for operation. Spectrum Allocation (SA) is closely related to the channel characteristics and operations of primary users. Furthermore, spectrum allocation is affected by the activities of other CR users in the network. The task of spectrum allocation consists of two parts, namely spectrum selection, and spectrum access. Spectrum allocation (or assignment) is responsible for assigning the most appropriate frequency band(s) at the interface(s) of a cognitive radio device according to some criteria (i.e., maximize throughput, fairness, spectral efficiency, etc.), while, at the same time, avoid causing interference to primary networks operating in the same geographical area.

First, each spectrum band is characterized, based on not only the local observations of CR users but also statistical information of primary networks, such as the average duration of their active or idle cycles, achieved through the sensing cycle. Then the most appropriate spectrum band should be selected, considering the SU's application requirements such as QoS, Bit Error Rate (BER), power limitation, and spectrum characteristics, for example availability of Channel State Information (CSI) for those channels at certain SUs, and the interference inflicted on those if the channels are occupied by other SUs.

The second part of spectrum allocation is spectrum access. Upon having selected the appropriate frequency band in the first stage, a CR should configure its transceiver parameters such as data rate, coding and modulation schemes, and power level for a proper access of the decided channel, considering the requirements of underlying application. In some applications, maintaining a high data rate is more important than packet loss criteria, therefore special modulation and coding schemes should be chosen to provide high spectral efficiency, whereas in other applications error performance can be more critical, which demands modulation and coding schemes with low bit error rate (BER) characteristics. Spectrum access techniques also take care of coordination among multiple secondary users in order to allow the fair use of spectrum and a smooth operation.

The procedure for solving the spectrum assignment problem in Cognitive Radio Networks (CRNs) is usually split in three steps. First, criteria (which define the target objectives) are selected to solve the SA problem. The second step includes the definition of an approach for modeling the SA problem that best fits to the target objective. The third and final step is the selection of the most suitable technique that will simplify and help solve the SA problem.

There are several criteria for assigning spectrum to SUs in cognitive radio networks, and these vary according to the target objectives of each algorithm. Interference is the most common criterion for designing an efficient cognitive SA algorithm. Cognitive radio networks have the constraint that the SUs should create no or limited interference to licensed users. Moreover, to maximize the CRN performance, interference between SUs should also be kept to a minimum.

A key objective for the deployment of cognitive radio networks is to achieve better utilization of the available spectrum bands. Thus, maximizing spectrum utilization is another common criterion for designing an efficient cognitive SA algorithm. Throughput maximization is also a very common criterion for channel assignment schemes in both traditional wireless networks and CRNs. The criterion of maximizing user or network throughput can create unfairness in the spectrum distribution among SUs, e.g. in cases where one SU can select multiple channels and others are left with no available spectrum (starvation). To avoid these cases, the criterion of maximizing throughput fairness among SUs have been considered [33], using several utility functions, e.g. by maximizing the minimum average throughput per SU, which leads to more fair results, or by using a fairness factor.

Spectrum assignment is also combined in several cases with another QoS criterion: the delay. End-to-end delay is the total time for the delivery of a packet measured from the source to the destination. In cognitive SA the end-to-end delay is considered in many approaches that combine spectrum allocation with routing. There are more criteria in SA such as Energy Efficiency, Network connectivity, price, and risk which are not explained here and interested readers may refer to [33].

TABLE 2.1. Criteria used for spectrum allocation.

Approach	Characteristics	Advantages	Disadvantages
Centralized	Spectrum server receives measurements from SUs and takes decisions.	Optimal decisions. Can achieve fairness between SUs. Can use priorities for most important SUs.	High signaling between SUs and the spectrum server. Not robust to spectrum server failures.
Distributed	SUs take decisions either as standalone or in cooperation with other SUs. No central entity exists.	Faster decisions. High flexibility - can adapt quickly to network conditions. Low signaling overload.	Not optimal decisions (only locally). Very difficult to achieve fairness among SUs.
Multi-Channel Selection	Spectrum aggregation. Capable of transmitting on multiple spectrum fragments with one radio interface.	Higher data rates. Maximum spectrum utilization.	Higher switching overhead. May increase interference when transmitting in multiple channels.
PU Considered	PUs’ presence is included in the decision. Target is to avoid interfering not only between SUs, but also with PUs.	More realistic approach.	Needs cooperation with PUs.
PU Not Considered	Only SUs are taken into account. The goal is to avoid interference between SUs and maximize their utilities.	Simplified approach.	Needs a predefined set of channels, which may become unavailable at a later time.
Segment-based	Network is divided into segments, the nodes of each have at least one common channel. Gateway nodes connect the segments.	Simple approach, achieves less channel switching.	Requires cooperation between nodes and an initial handshake between them. Gateway nodes could be easily congested.

When the SA problem has been defined based on specified criteria, the next step would be choosing an approach for solving the problem. Table 2.1 presents some basic approaches for spectrum assignment and summarizes their characteristics. More detail is available in [33].

And the last step in spectrum assignment process is selecting a proper technique for it. There are many techniques available for this task, and Table 2.2 summarizes some basic ones with their characteristics, advantages and disadvantages [33].

TABLE 2.2. Basic spectrum allocation approaches.

Approach	Characteristics	Advantages	Disadvantages
Heuristics	Iterative algorithms, finding at each iteration the best local solution , i.e. best available channel for SU.	Simplicity, easy implementation, speed, can be insensitive to specific problem characteristics.	Most developed approaches are problem-specific. There is no analytical methodology for studying their convergence. Can be limited to "local minimas" and not to optimum solutions.
Graph Theory	CRNs are visualized as graphs. Use conflict graphs, graph coloring or bipartite matching. Interference is modeled using conflict graphs.	Use existing solutions of graph theory.	Simplified assumptions. Cannot incorporate all parameters of CRNs, such as QoS requirements, ACI, etc.
Game Theory	SA is modeled as a game where the SUs are the players. Solution is found through Nash equilibrium. Can use several utility functions.	Solid decision making framework. Can be used for both cooperative or non-cooperative approaches	Difficult to structure the game in a way to guarantee equilibrium is always reached.
Linear Programming	The join power control/SA problem can be modeled as a MINLP problem and then into a BLP problem that contains only binary parameters.	Use of existing LP techniques.	Transformation from MINLP into BLP is not ensured and requires several assumptions, e.g. binary (0, max) transmission power.
Fuzzy Logic	Uses a set of rules for decisions, utility and membership functions for optimization and weights respectively.	Fast decisions based on the predefined rules. Learning techniques can improve the quality of the decisions.	Limited functionality because the rules are predefined. Needs a large number of rules to consider all parameters of CRNs. It is hard to determine accurate rules.

3) Spectrum Sharing:

Once a spectrum band has been selected through allocation process, spectrum sharing comes into action for accessing that band through adapting transmit power, bandwidth, and modulation scheme, constrained by power and the primary and secondary's error rate requirements.

There are two different sharing categories; sharing between the primary and secondary networks, and sharing among the CRN users. The underlay, overlay, and interweave paradigms explained in Section 2.1.1 are the answers to the first category problem.

Sharing the channel by CR users in the second category requires coordination of transmission attempts between them. There is possibility of multiple users competing to use the same portion of spectrum, which would result in collision among them. In this respect, spectrum sharing should incorporate an algorithm for accessing the spectrum which minimizes the risk of collision. This is inherently a function of the MAC layer. Power control and scheduling have some impact on functionality of higher layers too. For example, routing at transport layer is affected by the link capacity which itself is dependent on power control and scheduling. Constraints at application layer should also be considered while designing a power allocation scheme. Because some applications such as video transmission are power hungry whereas another group of applications, e.g., those involving sensor networks are power restricted. There is rich literature [34], [35], [36] available providing good insight on spectrum sharing fundamentals such as capacity and potential of spectrum sharing systems, and optimal limits of spectrum sharing.

Spectrum sharing techniques can be categorized into different classes based on the architecture of the underlying system, spectrum allocation behavior, and spectrum access techniques. Two classes of spectrum sharing techniques can be defined based on the architecture of the system, namely centralized spectrum sharing and distributed spectrum sharing. Centralized spectrum sharing employs a centralized entity to control the spectrum allocation and access activities of all the SU members, whereas distributed spectrum sharing does not need any governing entity and allows the CR users to handle spectrum allocation and access. Spectrum allocation method classifies the spectrum sharing techniques into two categories: cooperative spectrum sharing, and non-cooperative spectrum sharing. Cooperative spectrum sharing, as the name suggests, encourages participating SUs to share information with each other. Relay networks are one example of this sharing scheme. Non-cooperative spectrum sharing techniques do not rely on collaboration and consider only the node at hand.

4) Spectrum Mobility:

The CR environment is dynamic, and the frequency bands it is operating on may be reoccupied by the PUs at any instant. Therefore, a CR user also needs to be capable of changing its operating frequency when required, or randomly before the channel becomes unavailable. The process of SUs changing their operating frequencies can be defined as spectrum handoff, in which a SU should pause its current communication, vacate the channel, and determine a new available channel to continue its operation. When a CR needs to move to another spectrum to continue its transmission, spectrum allocation procedures should be followed to pick a new channel. As in spectrum allocation, spectrum mobility depends on sensing information, routing information, data link layer delay, hand-off delay, and the specified application.

These four CR-specific functions are additional to the main functions (such as coding/decoding, modulation/demodulation, routing, ...) required by any wired or wireless communication system. In a typical communication system, each function is implemented in a specific network layer. In a CR system, however, any of the four spectrum management functions incorporate some other functions and interactions inherent to more than a single layer. Therefore, in designing a CR network, cross-layer methods should be considered. Figure 2.1 illustrates spectrum management functions which span over different network layers.

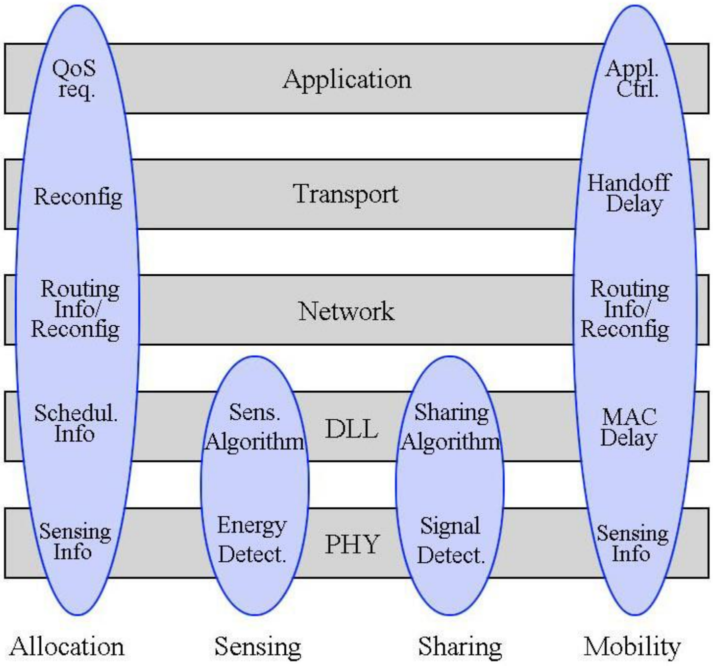


Fig. 2.1. Cognitive radio spectrum management framework; four main functions spanning over different OSI layers.

2.1.3 Applications of Cognitive Radio

Cognitive Radio is the proper candidate for many emerging applications and services running or relying on a wireless media. In this section we indicate some of them in brief.

Machine-to-Machine (M2M) technology:

Machine-to-Machine (M2M) technology and Internet of Things (IoT) have been around for quite some time and have rapidly grown over the last decade. In the near future M2M communication will be utilized in many applications such as industrial automation, logistics, transportation, smart grid, smart cities, healthcare, and defence, mostly for monitoring but also for control purposes [37]. It is envisaged that by 2020, over 50 billion devices will be connected to the internet, mostly wirelessly. The wireless interconnectivity of such a large number of devices has many implications. The more wireless nodes, the more bandwidth is required. Interference is another major issue; if all such devices are supposed to operate in the industrial, scientific and medical (ISM) bands it will result in high interference. Incorporating cognitive radio technology can mitigate both issues in such environments by employing smart techniques for accessing the wireless spectrum in an opportunistic manner while avoiding interference [37]. In [38] a cross-layer mechanism with joint power control, dynamic link scheduling and routing has been proposed for sensor networks used in M2M hazard monitoring applications, and in [39] a communication module for IoT is considered which is suitable for a cognitive radio platform.

Public safety networks:

Wireless communications are extensively used by emergency responders, such as police, fire and emergency medical services, to prevent or respond to incidents, and by citizens to quickly access emergency services. Video surveillance cameras and sensors are becoming important tools to extend the eyes and ears of public safety agencies. Correspondingly, data rates, reliability and delay requirements vary from service to service. Cognitive radio was identified as one of the emerging technologies to increase efficiency and effectiveness of spectrum usage in public safety networks. With cognitive radio, public safety users can use additional spectrum such as license-exempt TV white spectrums for daily operation from location to location and time to time. With appropriate spectrum-sharing partnerships with commercial operators, public safety workers can also access licensed spectrum and/or commercial networks.

Cellular networks*:*

The rapid proliferation of media-rich mobile devices has caused users to demand ubiquitous mobile broadband that can provide services comparable to the fixed broadband internet. To accommodate the exploding traffic, fifth generation (5G) cellular communication systems have been proposed by research communities worldwide. The challenges faced by 5G cellular networks are multifold and cognitive radio is a promising technology to tackle, or at least partly address, some of the challenges in 5G cellular networks [40]. CR technology is regarded as a promising solution to some of the challenges in the existing cellular networks. In certain geographical areas, cellular networks are overloaded, due partly to limited spectrum resources owned by the cellular operator. Many papers have investigated the application of spectrum sensing or spectrum sharing in cellular networks [41], [42] and have suggested how cognitive radio technology can augment high speed cellular networks such as LTE and WiMAX to dynamically use these newly available spectrums either in the access or backhaul parts of their networks. For access network applications, the large amount of data generated in a small area (such as an audience in a stadium or an arena), will drive both cellular networks and ISM band WiFi networks to become overloaded. This huge amount of data can be offloaded to additional spectrum, such as TV White Space (TVWS) by means of cognitive radio technology. In another scenario, with the use of cognitive radio technology, white space spectrum can be made available for cellular operators to use it for backhaul, to connect their cell towers to their backbone networks, thus reducing intensive backhaul cables installation, and provide coverage to more customers in unserved and underserved areas.

2.2 Full Duplex Communication

Conventionally, communication systems operate in half-duplex or out-of band full-duplex mode, meaning that they transmit and receive either at different times, or over different frequency bands. However, in In-Band Full Duplex (or briefly Full Duplex)[1] systems, data signals are transmitted and received simultaneously over the same frequency band. In the past, most researchers thought it is generally not possible for radios (e.g., base stations, relays, or mobiles) to receive and transmit on the same frequency band at the same time because doing so would result in strong self-

[1] Full Duplex and In-Band Full Duplex are used interchangeably in this thesis.

interference from the transmitter to the receiver which overwhelms the desired signal at the receiver. However, IFDB communication has a long history and has been used in the design of in-band wireless relays (where in-band repeaters are used to increase coverage in a wireless communication system by receiving, amplifying, and re-transmitting the wireless signal in the same frequency band) as well as short range continuous wave (CW) radar systems since the 1940s. But these applications were just special cases with many limitations, not suitable for general wireless systems.

Ideally, using the full duplex transmission mode means wireless communication systems can potentially double their spectral efficiency relative to systems with half-duplex (HD) operation. In addition to doubling the spectral efficiency, full duplex has other advantages which are explained below.

2.2.1 Advantages of Full Duplex Communication

- Capacity Enhancement: Full utilization of time and frequency resources make it theoretically possible for IBFD transmission to double the link capacity compared to HD transmission [6], [43].

- Feedback delay reduction: Reception of feedback signaling (such as control information, channel state information (CSI), acknowledge/nacknowledge (ACK/NACK) signals, resource allocation information, …) during data signal transmission, results in reduction of the transmission delay while enhancing the node throughput [44].

- End-to-end delay reduction: In relay systems, relay nodes with FD transmission can reduce end-to-end delay because the relay node simultaneously receives data from a source node and transmits data to a destination node [6], [45].

- Network fairness progress: The network fairness can be improved in many types of FD networks such as centralized networks. In HD centralized networks with n subscribers, all nodes including the access point (AP) get the same share of the channel, which is equal to 1/(n+1). However, since the data load of an AP is n-times more than a regular node, the channel fairness is not the same for all nodes. Using FD technology, the AP is able to transmit at all times concurrent with the transmission from other nodes. With full duplex, whenever the AP gets an upstream packet from any node, it is able to send a downstream packet to the same node, thus achieving fairness between upstream and downstream flows [46].

- Network security and secrecy improvement: An efficient way to increase the secrecy rate in wireless systems is to degrade the decoding capability of the eavesdroppers by introducing controlled interference, or artificial noise. The use of simultaneous transmission at two nodes means that eavesdroppers encounter a scrambled signal that is a blend of two different signals on a single channel, which cannot easily be separated with no side information [47].
- Lessening the hidden terminal problem in ad hoc networks: Because all nodes are transmitting, IBFD transmission can solve the 'hidden node' problem in ad hoc networks. Furthermore, the fact that simultaneous listening and sensing is being performed on a frequency band while the signals are being transmitted means that each node can decide whether or not the other nodes are transmitting signaling and thus prevent collisions [48].

2.2.2 Self-Interference and Cancellation methods

Despite this attractive list of advantages with IBFD transmission, it has not come into widespread use due to the severe effects of its inherent self-interference (SI). Because of the short distance between the transmit and receive antennas at a node (rather than the distance between two nodes), the SI power is much stronger than that of the desired signal. Self-interference refers to the interference that a transmitting IBFD terminal causes to itself, which interferes with the desired signal being received by that terminal (Fig. 2.2). For example, a MS at the edge of a small-cell can incur path losses of 110 dB. The maximum allowable transmit powers of the BS and MS are 24 dBm and

21 dBm [49], respectively, and the power of the received signal at the BS can be up to 21−111= −90 dBm. If we assume 15 dB isolation between the BS's transmit and receive signal paths, then the SI can be up to 24−15−(−90) =99 dB stronger than the desired received signal. If a reliable communication link requires a signal-to interference ratio (SIR) of at least 5 dB, then providing such a link would require a 104 dB reduction in SI [48]. As another example, in the worst cases in Wi-Fi systems110 dB of SI suppression is required to achieve ideal full duplex capacity if the SI channel does not have any path loss [7]. This strong SI can cause errors in hardware components in the receive RF chain, such as the low-noise amplifier (LNA), analog-to-digital converter (ADC), and mixer. This SI must be effectively canceled if IBFD systems are to become feasible.

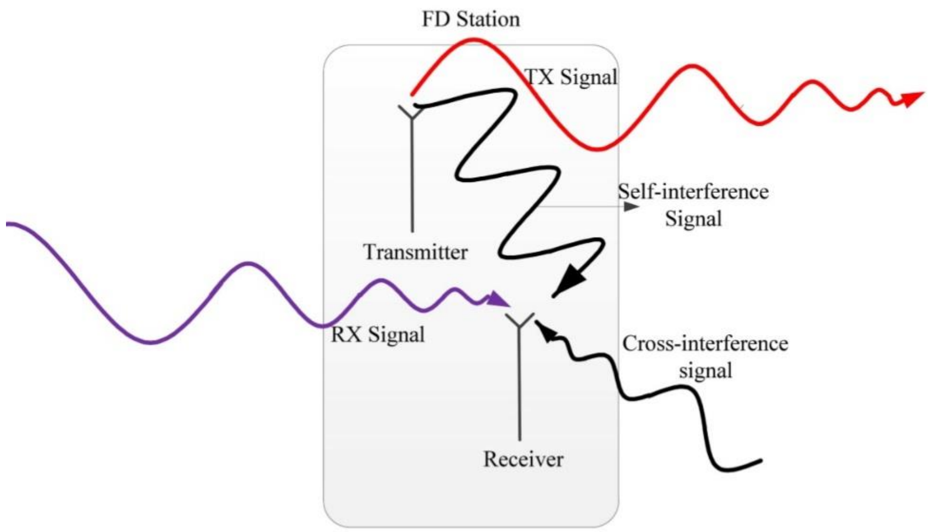


Fig. 2.2. Self-interference in a full duplex station.

In order to make IFDB possible, researchers have focused on methods to cancel self-interference. Existing self-interference mitigation technologies would also be divided into the following two categories: passive suppression, which electromagnetically isolates the transmit and receive antennas, and active cancellation, which exploits a node's knowledge of its own transmit signal to cancel the self-interference.

***Passive Self-IC Scheme*:**

Passive self-IC techniques suppress the SI before it enters into the receive RF chain circuit [49]. The amplitude of the received signal, therefore, can be reduced before the conversion process in the ADC in order to lighten the burden associated with supporting a wide dynamic range of received signals at the receiver. The most intuitive approach to passive cancellation is using path loss. The propagation path loss depends on the placement of antennas and the distance between the transmit and receive antennas. Typically, this is accomplished by some combination of antenna directionality, cross-polarization, and transmit beamforming. For IBFD relays, the earliest known self-interference suppression techniques relied on increasing the physical separation between transmit and receive antennas. Such methods are feasible when the system is equipped with separate transmit and receive antennas. However, in systems with shared antenna other methods are applied. The conventional CW radars of the 1940s and 1950s achieved isolation between the transmitter and receiver through antenna-separation-based path-loss in separate-antenna systems, or through the use of circulators in shared-antenna systems. A circulator is a passive non-reciprocal three- or four-port device, in which a microwave or radio frequency signal entering any port is transmitted to the next port in rotation. For a three-port circulator, a signal applied to port 1 only comes out of port 2; a signal applied to port 2 only comes out of port 3; a signal applied to port 3 only comes out of port 1. It is usually a Ferrite device which exploits the non-linear propagation of RF signals in magnetics materials (Fig. 2.3).

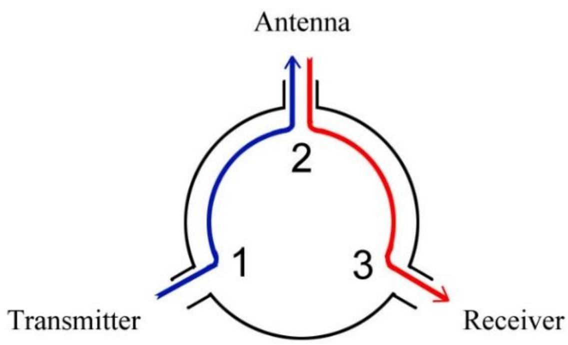


Figure 2.3. RF circulators used in shared antenna systems, like radars.

In radar, circulators are used as a type of duplexer, to route signals from the transmitter to the antenna and from the antenna to the receiver, without allowing signals to pass directly from transmitter to receiver. In a practical environment, however, the transmitted signal in the circulator flowsfromport1toport3, which can affect the signals received from port 2 as direct SI (the isolation in the circulator is less than 15 dB) [6].

It has been shown experimentally that up to 65 dB of SI can be suppressed with omni-directional (shared) antennas [48] and up to 72 dB with directional antennas [49].

Active Self-IC Scheme:

Analog cancellation is the first step in the active self-IC scheme. Its goal is to suppress SI on the analog circuit before it enters the ADC. Analog cancellation schemes are implemented by subtracting an estimated SI signal after the passive cancellation step. There are two types of analog cancellation [49]. The first uses the transmit signal to create a replica signal with a balanced/unbalanced (Balun) transformer [49] or with multiple parallel lines of varying delays and tunable attenuators [50]. The second method applies the attenuation and delay in the digital domain and converts it to the analog domain to cancel the SI [48]. Creating attenuation and delay is easier with digital processing than with analog processing. In [51], analog cancellation with 20 cm of antenna separation suppresses 72 dB of SI.

Digital cancellation is the last step in the active self-IC process and suppresses any residual SI after the ADC. Basically, the estimated residual self-interference is subtracted from the received signal in the digital domain [48]. There are a number of digital self-IC schemes using minimum mean square error (MMSE) receivers and multiple antennas, such as zero-forcing (ZF) beamforming, MMSE filter and null space projection (NSP) schemes [52].

By exploiting various combinations of the advanced self-IC techniques for the antenna, analog, and digital domains in an IBFD system, therefore, we can conclude that the power of the residual SI can be controlled, thus making IBFD transmission feasible.

2.3 Full Duplex Cognitive Radio

As explained in previous sections, both cognitive radio and full duplex communication are two promising technologies for enhancing spectrum utilization and when combined together in a network will improve network performance in terms of throughput and other critical metrics. In fact, FD is a capability and CR is a networking paradigm. FD communications may replace any HD communications in all wireless systems to some extent, but the CR idea may not be employed in all wireless platforms. Therefore, whenever we talk of full duplex cognitive radio, we mean a CR system in which FD operation is employed in some ways. There are extensive studies on CR scheme in HD wireless systems, which are all candidates for being updated for operation in FD mode [53], [54].

As mentioned above, adding FD capability to a CR network will result in higher throughput achievements for the secondary network. But this is not the only motivation for application of FD operation in CR networks. Apart from other advantages of FD systems mentioned in 2.2.1, the cognitive systems may benefit from FD operation in a special way. All cognitive systems need to sense the spectrum before and during transmission to avoid interference on the incumbent users. In HD systems it is not possible to sense and transmit at the same time, which results in degradation of throughput and the primary's protection. With FD capability, the secondary users can sense and transmit at the same time, which will enhance the whole system's performance considerably, as explained in the following chapters.

As a new field of research, there are not many studies on full duplex cognitive radio networks. Most studies on full duplex cognitive radio have concentrated on simultaneous transmitting and sensing (i.e. full duplex spectrum sensing) in the interweave paradigm [55]-[59]. Traditional spectrum sensing schemes, such as Listen Before Talk (LBT) in non-time slotted networks, usually do not guarantee good protection for the PUs, as they may become active at any time during the SUs' transmission, and this will cause collision. Some studies [55]-[58] have proposed full duplex spectrum sensing to alleviate this problem. In such works a transmitting SU keeps sensing the presence of the PU constantly during the transmission slot, and upon detection of a PU signal, will stop transmission to avoid collision [55]. The same idea has been studied in [60] a network with directional multi-reconfigurable antennas.

Tsakalaki et al. then described a structure to enable simultaneous communication and sensing using spatial filtering with a three-port antenna system (two antennas for transmission and one antenna for sensing) [61]. Expanding on those works to the scenario with multiple antennas, Heo et al. proposed a concept of "TranSensing" which adaptively uses spatial resources for simultaneous PU sensing and SU data transmission, and then investigated the information-theoretic performance of the proposed scheme [62].

In another group of studies, the notion of simultaneous transmission and reception of a node has been investigated. [63], [64] proposed the simultaneous "transmission-and-reception (TR) mode" as a way of increasing the performance of full-duplex-based CR networks in terms of both spectral efficiency and outage probability. More specifically, optimal SU power allocation and a mode selection algorithm based on the residual SI value were developed to decrease the SU outage probabilities by taking residual SI into consideration. When the residual SI is large, each node operates in HD transmission mode rather than in IBFD transmission mode based on the theoretical analysis in [63].

A similar concept has been developed by Afifi et al. in [65]. They proposed a similar approach for a FD scenario in a CR network with four modes of operation; sensing-only, transmit and sensing, transmit and receive, and channel selection, and developed an optimal mode-selection strategy to maximize the SU throughput for a given PU collision probability. The criterion for choosing the optimal action was to maximize the SU's utility subject to a constraint on the PU collision probability. They considered waveform-based spectrum sensing technique for the transmit and sensing mode, in order to overcome detection inaccuracies at imperfect SIS inherent to energy detection methods.

Some studies have focused on enhancing cooperative schemes in CR networks with the advantages of FD communications. In HD systems, a cooperating node may pause its own job, and take part in relaying another network's packets or in some scenarios cooperate in the sensing procedure. When FD is feasible, the cooperating node may assist other networks at the same time it is serving its own clients. Listen and Talk is an example of such ideas proposed by [67] in a cooperative CR network, where secondary nodes take part in cooperating sensing at the time of transmission. Authors in [68] considered the cooperation between a primary system and a cognitive system in a cellular network where the cognitive base station (CBS) relays the primary signal using amplify-and-forward (AF) or decode-and-forward (DF) protocols, and in return it can transmit its own cognitive signal.

As mentioned earlier, FD transmissions may be implemented in almost every CR network design based on HD operation, in order to achieve improved performance of the overall network. In the following chapters of this thesis we have investigated such implementation in two existing CR schemes and demonstrated the respective advantages of FDCR technology.

Chapter 3

Cooperative ARQ in Full Duplex Cognitive Radio Networks

3.1 Introduction

In a CR network, the secondary users are allowed to share the licensed spectrum of a primary network as long as the interference impinged on the primary network is within tolerable limits. To accomplish this goal, a cognitive user monitors the ambient licensed or unlicensed radio spectrums and try to identify any opportunity arising during the primary networks' operation suitable for transmission, and dynamically adapts its operation parameters in a way to utilize them efficiently. These opportunities may be in the form of unused spectrum bands (spectrum holes) in the time or frequency domain in the interweave paradigm, or variable conditions of the channel between primary source node to primary destination node, which enables SUs to have concurrent transmissions in underlay or overlay schemes.

In the interweave network paradigm, secondary users (SUs), identify the unused slots of the primary users' licensed spectrum in time or frequency domain and will transmit their data as long as the primary signal has not reappeared. [70], [71] are some studies on this type of paradigm in half duplex networks. Such methods are based on zero interference/collision rationale which may be achieved through accurate spectrum sensing protocols. On the other hand, in underlay and overlay paradigms, there are many different types of transmission opportunities for CR users. In these schemes, a cognitive user may transmit at the same time and in the same frequency band of the primary network, as long as the interference caused by the secondary on the primary's signal is within tolerable limits. To meet these criteria, a secondary network would adapt its transmission parameters such as rate, power, and coding, based on the primary network's conditions and the respective opportunities. Such opportunities are usually related to the status of the primary's channel (i.e. between the primary transmitter and the primary receiver). When this channel is in a very good condition, a power-controlled SU transmission will not cause considerable interference on the primary's signal and would be easily tolerated. On the other hand, a deeply faded channel which is not capable of successful transmission of primary's data, implies another opportunity for the secondary network to transmit freely with no concern about inflicting interference on the primary's signal. Sometimes an SU may create such an opportunity for itself through cooperation with a PU. In such scenarios, an SU such as a relay node forwards the primary's signal to the primary destination node by cooperation, and in turn the SU system is allowed to utilize the primary spectrum as a reward. This concept has been proposed in HD based CR networks under the name spectrum leasing concept [43]. In overlay schemes where the secondary network knows the primary's codebook and transmission parameters, the secondary users may employ methods such as network coding to create opportunities for concurrent data transmission.

In order to use such opportunities in an underlay paradigm, the secondary network requires having perfect channel state information (CSI) describing the primary's channel. Such information may be acquired through aggressive or non-aggressive methods. In a non-aggressive method, the secondary network should have *a priori* information about the primary's transmission parameters and should be capable of gathering CSI through the eavesdropping primary's feedback channel, which is not always possible. In aggressive methods, secondary users gather such information through probing mechanisms. The SU will transmit a probe packet occasionally and through overhearing the primary's feedback signals (ACK/NACK) will be informed to some extent of the condition of the primary's channel [73]. An acknowledgement (ACK) signal means that the channel is in a good state despite the presence of the probe signal, whereas a NACK may imply that the channel is not good enough to tolerate any additional interferences. As the name implies, such methods are aggressive and cause interference and degradation to the primary's performance corresponding to the frequency of probing.

However, some opportunities are not directly related to the condition of the primary's channel, and partial information about the primary's transmission status, achieved through non-aggressive methods, enables the secondary network to enter into action. One of such opportunities arises during primary network ARQ retransmissions. The replicas of the primary data retransmitted during ARQ rounds create redundancy in the primary channel and the SU can take advantage of it through interference cancellation [77]. In other words, the first transmission has already provided some information to the primary receiver (albeit not enough), the second transmission now needs to provide less than full amount of mutual information and can be more robust to interference, and this condition creates an opportunity for a secondary network to transmit some data. If these opportunities are properly exploited, nontrivial rates can be harvested for secondary users with little or no degradation to the primary's performance. In this chapter we have proposed a scheme which exploits these opportunities for secondary network communication. The scheme presented is based on the FD operation of one the secondary users during the ARQ round of the primary network, and incorporates exploitation of the Alamouti scheme and cooperation of the primary and secondary networks. Packet error rate(PER) and throughput of both networks have been analyzed as the performance metrics of the system. In this section we describe some basic elements considered in this chapter.

Automatic Repeat reQuest (ARQ)

The Automatic Repeat reQuest, also known as Automatic Repeat Query is a control mechanism of data link layer where the receiver asks the transmitter to send again a block of data (packet or frame) when errors are detected. The ARQ mechanism is based on acknowledgement (ACK) - messages sent by the receiver indicating that it has correctly received a data frame or packet - and timeouts (specified periods of time allowed to elapse before an acknowledgment is to be received) to achieve reliable data transmission over an unreliable service. If the sender does not receive an acknowledgment before the timeout, it usually re-transmits the packet/frame until the sender receives an acknowledgment or exceeds a predefined number of re-transmissions. In some systems nonacknowledgement (NACK) messages are used instead of ACK messages, indicating that the receiver has not received the previous block of data correctly. Then the transmitter will retransmit the previous block of data upon reception of a NACK message.

There are three types of ARQ protocols, *Stop-and-wait ARQ*, *Go-Back-N ARQ*, and *Selective Repeat ARQ / Selective Reject* [78].

Stop-and-wait ARQ is the simplest ARQ. It has one block of data at a time sent with no additional blocks sent until reception of the previous one is confirmed via an acknowledgement signal.

Go-Back-N ARQ is a more complex protocol. It allows blocks to be sent even if previous blocks were received without an acknowledgement signal. This protocol keeps track of the sequence. When the last block is received, it requests re-transmission of the blocks sent without an acknowledgement. This is repeated until all blocks are received with an acknowledgement signal. However, this protocol may result in many blocks being sent multiple times, which can be avoided by using the Selective Repeat ARQ protocol.

Selective Repeat ARQ may be used for the delivery and acknowledgement of sent data packets or the delivery of subdivided messages in sub-units. In the first methodology, the protocol continues to accept and acknowledge frames sent after an initial error. It continues doing this until a specified number of frames have been received, called window size. There is a window size for both transmitting and receiving and they must be equal. The sequence numbers of all frames not received are tracked and sent back to the transmitter.

ARQ protocols reside in the Data Link or Transport layers of the OSI model. TCP uses a variant of Go-Back-N ARQ to ensure reliable data transmission over the internet protocol. However, it does not guarantee delivery of data packets. When local area networks (LAN) have noisy environments, Selective Repeat ARQ is employed with packet segmentation.

Hybrid ARQ

The Hybrid ARQ (HARQ) mechanism uses an error control code in addition to the retransmission scheme to ensure a more reliable transmission of data packets (relative to ARQ). The main difference between an ARQ scheme and an HARQ scheme is that in HARQ, incorrectly received coded data blocks are often stored at the receiver rather than discarded, and when the re-transmitted block is received, the two blocks are combined in order to improve reliability. While it is possible that two given transmissions cannot be independently decoded without error, it may happen that the combination of the previously erroneously received transmissions gives us enough information to correctly decode. HARQ parameters are specified and negotiated during the initialization procedure.

There are two main combining methods in HARQ [78]:

- ***Chase combining***: every re-transmission contains the same information (data and parity bits). The receiver uses maximum-ratio combining to combine the received bits with the same bits from previous transmissions. Because all transmissions are identical, Chase combining can be seen as additional repetition coding. One could think of every re-transmission as adding extra energy to the received transmission through an increased Eb/N0.

- ***Incremental redundancy***: every re-transmission contains different information than the previous one. Multiple sets of coded bits are generated, each representing the same set of information bits. The re-transmission typically uses a different set of coded bits than the previous transmission, with different redundancy versions generated by puncturing the encoder output. Thus, at every re-transmission the receiver gains extra information.

Several variants of the two main methods exist. For example, in partial Chase combining only a subset of the bits in the original transmission are re-transmitted. In partial incremental redundancy, the systematic bits are always included so that each re-transmission is self-decodable.

HARQ is used in High Speed Downlink Packet Access (HSDPA) and High Speed Uplink Packet Access (HSUPA) which provide high speed data transmission (on downlink and uplink, respectively) for mobile phone networks such as Universal Mobile Telecommunications Service (UMTS), and in the IEEE 802.16-2005 standard for mobile broadband wireless access, also known as "mobile WiMAX". It is also used in Evolution-Data Optimized and LTE wireless networks.

Alamouti Scheme

The Alamouti scheme is the simplest of all the space-time block codes (STBC) which are used for MIMO systems to enable the transmission of multiple copies of a data stream across a number of antennas and to exploit the various received versions of the data to improve the reliability of data-transfer. Space-time coding combines all the copies of the received signal in an optimal way to extract as much information from each of them as possible. In basic Alamouti scheme there are two transmit antennas and one receive antenna. This technique enables the system to achieve transmit diversity of order two despite the fact that channel knowledge is not available at the transmitter. It works as follows [84]:

At a given symbol period, two signals are simultaneously transmitted from the two transmitters (antennas) as shown in Fig. 3.1. The signal transmitted from transmitter 1 is denoted by x_1 and the one from transmitter 2 two by x_2. During the next symbol period signal $-x_2^*$ is transmitted from transmitter 1, and signal x_1^* is transmitted from transmitter 2 where * is the complex conjugate operation. h_1 and h_2 are complex path gains from transmitter 1 and transmitter 2 to the receiver, respectively. This sequence is shown in Table 3.1.

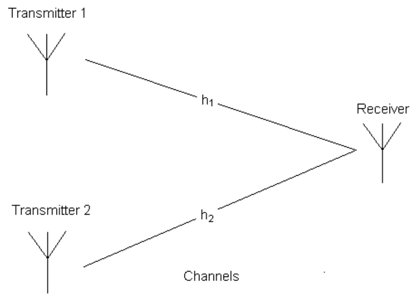


Fig. 3.1. Alamouti scheme transmission diagram

Table 3.1. The encoding and transmission sequence of Alamouti scheme

	Transmitter 1	Transmitter 2
Time t	x_1	x_2
Time t+T	$-x_2^*$	x_1^*

The received signals can then be expressed as

$$r_1 = r(t) = h_1 x_1 + h_2 x_2 + n_1$$
$$r_2 = r(t+T) = -h_1 x_2^* + h_2 x_1^* + n_2 \qquad 3.1$$

where r_1 and r_2 are the received signals at time *t* and *t*+*T*, and n_1 and n_2 are complex random variables representing receiver noise and interference at the receiver. The two signals are combined at the receiver and the following two signals are built:

$$\tilde{x}_1 = h_1^* r_1 + h_2 r_2^*$$
$$\tilde{x}_2 = h_2^* r_1 - h_1 r_2^* \qquad 3.2$$

These two combined signals are then sent to the maximum likelihood detector which, for each of the signals x_1 and x_2, applies the decision rule used for PSK signals. The resulting combined signals in (3.2) are equivalent to that obtained from conventional two-branch MRRC. The only difference is phase rotations on the noise components which do not degrade the effective SNR. Therefore, the resulting diversity order from the Alamouti scheme with one receiver is equal to that of two-branch MRRC with one transmitter and two receivers.

Packet Error Rate (PER)

The packet error rate is an important quality of service parameter for wireless networks. It is the number of incorrectly received data packets divided by the total number of received packets. A packet is declared incorrect if at least one bit is erroneous.

PER is mostly derived using BER and an error independence assumption. In general, having the instantaneous received SNR γ at the receiver, we can calculate the exact physical layer packet error rate of a simple packet-based transmission system if FEC is not used as follows [85]:

$$PER(\gamma) = 1 - (1 - BER(\gamma))^n, \tag{3.3}$$

where $BER(\gamma)$ is the bit error rate, γ is the received SNR, and n is the length of the packet. If a block code with k-bits error correction capability is used, the $PER(\gamma)$ can be written as [85]:

$$PER(\gamma) = 1 - \sum_{i=0}^{n} \binom{n}{i} (BER(\gamma))^i (1 - BER(\gamma))^{n-i}, \tag{3.4}$$

As we see the exact formulation of PER is very complicated for analysis, especially when we have large-sized packets (in practice packets are of sizes 1080 bits or more according to the wireless standards such as HYPERLAN [86], and IEEE 802.22 [5]). Therefore, we seek to find a good approximation for our analyses. Q. Liu et al. proposed an approximation for packet error rate in [87] which is widely considered in many works and we use it here as well:

$$PER(\gamma) = \begin{cases} 1, & 0 < \gamma < \gamma_t \\ \alpha \exp(-g\gamma), & \gamma \geq \gamma_t \end{cases}, \tag{3.5}$$

where γ is the instantaneous signal-to-noise ratio and $(\alpha, g,\ \gamma_t)$ are mode dependent parameters found by least-squares fitting to the exact PER. For a packet length 1080 bits, the corresponding values for $\alpha, g,\ \gamma_t$ are provided in [87] which have been reproduced here in Tables 3.2 and 3.3. Table 3.2 is for un-coded modulation and Table 3.3 for convolutionally coded modulation.

TABLE 3.2. PER parameters for un-coded MQAM modulation [87]

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6	Mode 7
Modulation	BPSK	QPSK	8-QAM	16-QAM	32-QAM	64-QAM	128-QAM
Rate(bits/sym.)	1	2	3	4	5	6	7
a	67.7328	73.8279	58.7332	55.9137	50.0552	42.5594	40.2559
g	0.9819	0.4945	0.1641	0.0989	0.0381	0.0235	0.0094
γ_t (dB)	6.3281	9.3945	13.9470	16.0938	20.1103	22.0340	25.9677

TABLE 3.3. PER parameters for coded MQAM modulation [87]

	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5	Mode 6
Modulation	BPSK	QPSK	QPSK	16-QAM	16-QAM	64-QAM
Coding rate R_c	1/2	1/2	3/4	9/16	3/4	3/4
Rate (bits/sym.)	0.50	1.00	1.50	2.25	3.00	4.50
a	274.7229	90.2514	67.6181	50.1222	53.3987	35.3508
g	7.9932	3.4998	1.6883	0.6644	0.3756	0.0900
γ_t (dB)	-1.5331	1.0942	3.9722	7.7021	10.2488	15.9784

The switching threshold γ_t is set such that $\alpha \exp(-g\gamma_t) = 1$.

3.2 Literature review

The idea of exploiting the primary ARQ has been investigated in some works. In [74] and [75], a secondary system is allowed to transmit in the same frequency band provided it ensures that the primary attains a specified target rate. To accomplish this goal, they proposed a scheme in which the secondary eavesdrops on the primary's ARQ to estimate the throughput loss of the primary user and uses this knowledge to tune the transmission policy accordingly and stay within its interference budget. The cognitive radio transmitter and receiver listen to the primary’s ARQ, and by calculating the frequency of the primary’s ARQ, can estimate the rate of the primary network. Whenever the primary’s rate is less than the target rate, the secondary network will decrease its rate to cause less interference, and whenever the primary's rate is higher than target rate, the secondary may increase its transmission rate. In this way the cognitive radio can adjust its rate appropriately. Under certain assumptions, they showed there exists an optimal Rate-Interference Budget (RIB) tradeoff.

Similarly, in [76] the secondary user co-exists with a maximal transmission primary system, under a constraint on the maximum throughput loss of the primary. The secondary source aims to maximize its own throughput, while guaranteeing a bounded performance loss for the primary source. They derived the optimal transmission policy for the secondary user when the primary user adopts a retransmission based error control scheme.

[77] proposed a system in which an SU is allowed to transmit with the PU simultaneously during the primary retransmission period. Therein, the PU employs Hybrid ARQ with incremental redundancy, and a secondary pair is allowed to transmit its own data during the primary's retransmission rounds. This requires that the SU receiver be capable of cancelling the primary's interference. Therefore, the secondary network will transmit on the condition that it had acquired the primary's message correctly during its initial transmission to be able to cancel the primary's interference. In the proposed protocol in [77], the primary has one round of retransmission, but the secondary has no such opportunity. They calculated the outage probability for the primary and secondary pairs, and the respective achievable throughput. [79] and [80] developed the same idea when the SU is also allowed to retransmit some of its failed packets according to an optimal algorithm. They have exploited the temporal redundancy in the channel introduced by primary and secondary retransmissions for interference cancellation. The authors in [81] proposed a scenario wherein the secondary source is allowed to superimpose its transmissions over those of the primary source while guaranteeing a bounded performance loss for the primary network. This protection is achieved through monitoring the frequency spectrum of the primary's retransmission attempts and by subsequent adaptation of the secondary's transmission.

The work in [73] considers the same model as in [77] in a slow fading channel where the channel gain is assumed to be approximately constant over several transmission intervals, but is subject to change over much larger time scales. The authors proposed a method for acquiring partial information about the relative strengths of the cross channel and primary channel through listening to the primary ARQ feedback. The secondary transmitter also probed the channel by transmitting a packet and listening to the primary control signals (ACK/NACK), thus getting additional partial information about the relative strength of the cross channel and primary channel. Therefore, the secondary pair finds more chances for transmission and achieves higher throughput compared to that of [77]. They titled their method Spectrum Harvesting with ARQ Retransmission and Probing (SHARP) and proposed two varieties of it; conservative and aggressive SHARP. Conservative SHARP leaves the primary operations altogether unaffected, while aggressive SHARP may occasionally force the primary to use two instead of one transmission cycle for a packet, in order to harvest a better throughput for the secondary. Although their method did not introduce any outage in the primary, it would affect the primary's throughput.

All the above proposed methods impinge some degradation to the primary's throughput as a result of the secondary's interference. In order to mitigate this drawback, [82] utilizes cooperative ARQ and network coding. In their proposed protocol, the SU relays the primary packet and serves its own traffic simultaneously through a network coding scheme in the ARQ rounds. This requires coordination between primary and secondary networks for cooperative ARQ process. Usually, primary networks are oblivious to other networks' operation and this method may not be applied in many CR network scenarios. Although this method protects the primary network from additional interference at the cost of complexity, lower throughputs for secondary network are achieved.

In this chapter we propose a novel cooperative full duplex protocol for retransmission of the primary's ARQ packets based on the model considered in [77]. In contrast with other cooperative schemes in CR networks, which require coordination and synchronization of primary and secondary networks, in our proposed method the primary network is oblivious to the secondary's presence, but will benefit from its cooperation whenever the secondary network takes part. Thanks to the full duplex technology exploited in our scheme, non-trivial rates for secondary users are made available during the primary's ARQ rounds without any degradation to primary's performance in terms of packet error rate or throughput. The use of the Alamouti code in the cooperative retransmission of the primary packets is another key factor in our proposed protocol. Packet error rate (PER) of the primary and secondary networks for the proposed FD scheme and HD scheme proposed in [77] have been analyzed for the first time, which is a considerable aspect of this work.

The main contributions of this chapter are summarized as follows. We have proposed a novel cooperative protocol based on FD capability of secondary nodes to utilize the opportunities arising during primary ARQ retransmission intervals for SU's transmission. We have analyzed the packet error rate and throughput of the primary and secondary networks mathematically for ARQ and HARQ schemes and have shown that this protocol not only does not harm the throughput of the primary, but also will enhance it in the HARQ scheme due to incorporating a cooperative scheme. In contrast to other cooperative methods based on network coding, this method is easy to implement and higher throughputs are achievable for both primary and secondary systems. In addition, in this scheme PUs are oblivious to SUs and will benefit from the secondary's cooperation whenever applicable.

In the next section we describe the system model and our proposed protocol, and in Section 3.4 the packet error rate and throughput analyses in ARQ and HARQ scenarios are presented. We verify the validity of the proposed protocols by extensive simulations and present numerical results and the respective discussions in Section 3.5. We have summarized the used major notations in Table 3.4.

TABLE 3.4. Significant Notations.

P_1, P_2, S_1, S_2	Primary and secondary nodes
β	SIS factor
P_p, P_s	Primary and secondary transmission power, resp.
R_{p_0}, R_{s_0}, R_p, R_s	Primary and secondary nominal rates, actual rates, resp.
α, g, γ_t	PER fitting parameters
γ_I, γ_R	Initial and repeated signals SNR, resp.
$\gamma_{P_1,P_2}, \gamma_{P_1,S_2}$	P_1- P_2 and P_1- S_2 links signal-to-noise ratio, resp.
$\gamma_A, \gamma_B, \gamma_C, \gamma_D, \gamma_E$	P_1- P_2, S_2- P_2, S_1- P_2, S_2- S_2, & P_1- S_2 links SNR, resp.
$\mathcal{E}_T, \mathcal{E}_C, \mathcal{E}_{CH}$	PER of traditional, cooperative ARQ, and cooperative HARQ transmissions, resp.
P_P^e, P_S^e	Overall PER of primary and secondary networks, resp.

3.3 System Model and Protocol Description

As depicted in Fig. 3.2, we consider an overlay cognitive radio system with a pair of primary nodes (P_1 and P_2) and a pair of secondary nodes (S_1 and S_2). In each network, node 1 is the source node and node 2 is destination, i.e. data are transmitted from node 1 to node 2. In the secondary network, node S_2 operates in full duplex mode that is to transmit and receive at the same time over the same frequency band. To accomplish this purpose, node S_2 is assumed to be capable of perfect or partial self-interference suppression. There are many passive and active methods for self-interference

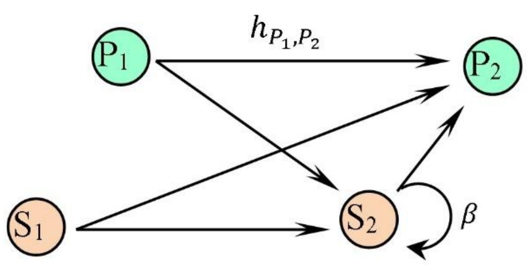


Fig. 3.2. System model of the network.

suppression as indicated in Chapter 2, and here we do not consider these methods. In the rest of this thesis, we just quantify this capability by SIS factor $\beta = \frac{\chi'}{\chi}$ which is the ratio between the residual self-interference signal (χ') and the original one (χ). In perfect SIS (no residual self-interference, $\chi' = 0$) β is zero; and otherwise, only the fraction $1 - \beta$ of the original self-interference signal is cancelled. As an example, for a residual self-interference signal of power (in watts) equal to 1% of the power of the original self-interference signal, this translates into $\beta = \sqrt{0.01} = 0.1$.

In addition, S_2 is capable of perfect Known Interference Cancellation (KIC). As the name implies, known interference in wireless networks occurs when the original data of the interference is known a priori. In theory, when a KIC-enabled node overhears and caches a packet, it becomes effectively immune to the interference caused by the subsequent transmission of that packet [83]. To accomplish this, the receiver first estimates the channel coefficient of the interference signal and then removes the known interference from the target signal. However, the scheme does not perform well when the channel estimation is inaccurate. By contrast, the blind known interference cancellation (BKIC) schemes proposed in [83] can obtain near-perfect performance and completely remove the interference while avoiding estimation of the interference channel. BKIC uses the interference in one symbol to cancel the interference in its adjacent symbol. For example, if the interference channel is h and the interference symbol is 1, then the interference in the current symbol is h. If the interference data in the adjacent symbol is -1, then the corresponding known interference is approximately –h. The interference can be cancelled with each other if we combine the two symbols. In this method, it is assumed that the secondary pair knows the codebook used by the primary transmitter and can decode the received primary signal which is our assumption as well, thus the primary's interference is known to S_2 and can be cancelled completely thanks to BKIC method.

All channels between nodes are assumed to be flat, quasi-static Rayleigh slow fading where the channel gains are assumed to be approximately constant over the original and retransmission intervals. The channel power gains between each of the two nodes, denoted by $h_{a,b}$ (where a and $b \in \{\mathrm{P}_1, \mathrm{P}_2, \mathrm{S}_1, \mathrm{S}_2\}$ represent the source and destination nodes respectively), follow an exponential distribution and the additive noise is assumed Normally distributed as $\mathcal{N}(0,1)$. It is assumed that the primary user P_1 transmits at a fixed power P_p, fixed nominal rate R_{p_0}, and employs truncated ARQ to reduce delays and buffer sizes. In this thesis, the maximum number of retransmission attempts is set to one. That is, a packet will be discarded if it is still incorrectly received after one round of retransmission. This assumption is quite suitable for real-time traffic, which tolerates a certain packet loss rate but requires very small packet delays. However, the method and analysis presented may be easily extended to systems with more than one retransmission. In this chapter we consider both ARQ and Hybrid-II ARQ (incremental redundancy which combines the signals of the original and repeated packets) concepts for the primary network and investigate the system performance for both scenarios. We assume that node S_2 cooperates with P_1 in retransmission of failed packets whenever applicable as explained in the next section. Primary nodes are oblivious to the presence of secondary nodes and the secondary users will detect the start and the end of the primary's transmission slot, which enables them to synchronize themselves with the primary network operation. In this way no special synchronization or coordination is required for the cooperative ARQ process which is one of the interesting features of the method proposed. S_2 has the same power and rate settings as P_1 (i.e. fixed power P_p and fixed rate R_{p_0}) when cooperating with P_1 in retransmission of failed packets, but S_1 transmits the cognitive messages at different settings (P_s, R_{s_0}) which may be fixed or variable. In this chapter we assume that secondary network does not generate ARQ feedback, however the analysis for ARQ enabled secondary network is straightforward.

We assume all codewords are drawn from a Gaussian codebook. The codewords fit within a coherence interval but are sufficiently large to allow reliable decoding whenever the mutual information of the channel supports the attempted rate. Further we assume that the secondary transmitter has no information about the primary's operation parameters or channel status (CSI) except overhearing the ACK (NACK) signals transmitted by the primary receiver at the end of the initial round of transmission. These feedback signals (ACK/NACK) are assumed to be error free, however we will consider the effect of non-perfect feedback on the performance metrics of the system briefly. The protocol works as follows.

3.3.1 Protocol Description

1) During the original transmission round, the primary node P_1 transmits a packet and both P_2 and S_2 attempt to receive and decode it. If P_2 receives and decodes this packet successfully, it sends an ACK which will be heard by P_1 and S_2. S_2 will discard the received signal from its buffer and P_1 gets ready for transmission of the next packet in the subsequent slot.

2) If P_2 cannot decode the packet, it will not send an ACK in the due feedback time slot. P_1 which is listening to the feedback channel not receiving any positive feedback, will imply that the packet was not received properly and will get ready for retransmission of the failed packet in the next transmission slot. On the other hand, S_2 which is monitoring the primary's transmission and feedback either (i.e. S_2 does not receive P_2's ACK within the feedback slot - with a known duration - upon termination of the primary's packet transmission), will come into action to take advantage of this opportunity provided that it has decoded the primary's signal successfully. If so, it declares a Clear to Send (CTS) signal to S_1, and will prepare for a full duplex operation in the ARQ round. It is assumed that CTS alarms are of shorter duration than ACK signals, and S_2 will have enough time to notice the lack of ACK and then emit the CTS alarm, all within the feedback time-slot. Or we may assume that the secondary network uses a report channel for emitting its own control signals such as CTS or ACK/NACK, to avoid collision with the primary's ACK signals.

3) In the ARQ round which starts immediately upon termination of the feedback (ACK) time-slot, P_1 retransmits the failed packet using STBC (Alamouti) [84] code regardless of S_2 cooperation. If S_2 had issued a CTS signal in the previous round, and it was heard by S1 during the primary's feedback slot (which will be the case as we assume CTS signals are error-free as well), then S_1 transmits its cognitive message (at power P_s) in the ARQ round, and S_2 will cooperate with P_1 in retransmitting the same packet (via Alamouti code) at the end of primary's feedback slot (which will synchronize both transmissions from P_1 and S_2), while receiving the cognitive message simultaneously in a full duplex manner. Otherwise, only P_1 will repeat the failed packet and both secondary nodes will remain silent. In the former case, S_2 can fully peel off the primary interference due to perfect known interference cancellation (KIC) capability, but the fraction β of its own signal (self-interference) will remain as a result of partial SIS capability.

As we have assumed the maximum of one retransmission for the primary and no ARQ for the secondary, each receiver will drop the respective packet upon failure in the second round, and transmission of a new packet would start in the subsequent slot. As both primary and secondary networks are aware of the exact duration of the feedback time-slot followed by each primary's transmission, their cooperative transmission will be well-synchronized without any coordination required. To accomplish this, P_1 and S_2 (as well as S_1) will start their transmission exactly at the end of ACK/CTS time-slot. The secondary users know the exact duration of the primary's packet transmission and feedback slot *a priori*, so they can synchronize themselves with the primary network through exact detection of the start and the end of P_1 packet transmission. Such detection would be feasible through any of the sensing methods of a cognitive radio network such as energy detection, cyclostationary detection, … explained in previous chapter. We have assumed that CTS signals are slightly shorter than ACK alarms in order that S_2 has sufficient time to decode a lack of ACK and then emit a CTS if applicable.

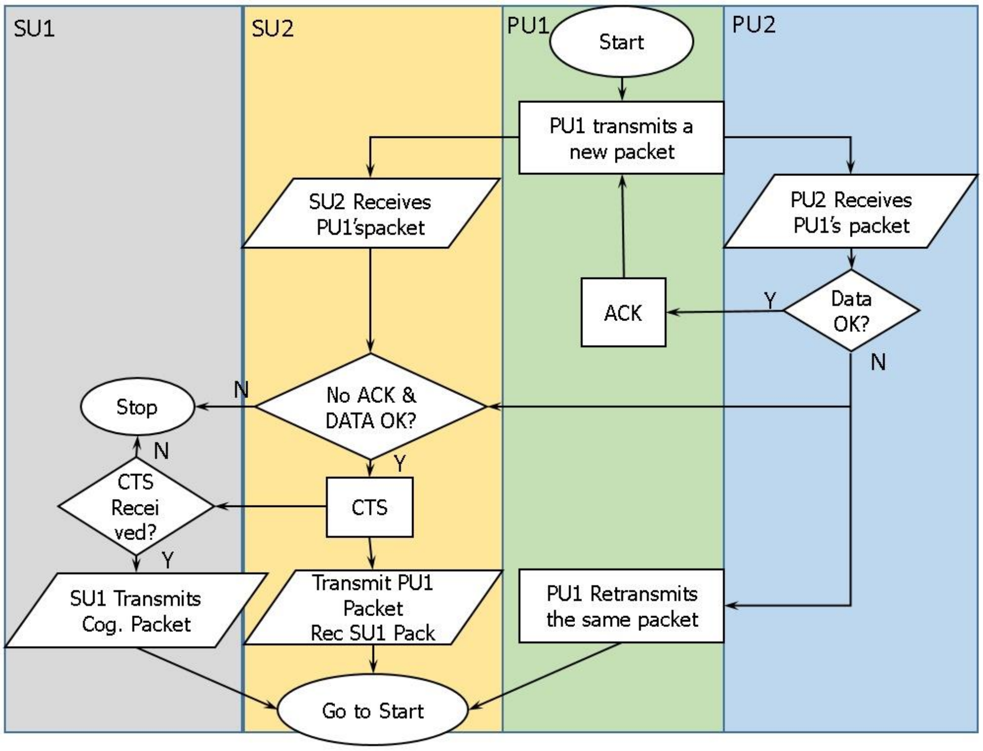


Fig. 3.3. Protocol flowchart.

The whole process is depicted in the flowchart shown in Figure 3.3.

The Alamouti scheme has a big role in the above protocol.

In our system, P_1 is transmitter 1, and S_2 is transmitter 2. P_1 always retransmits the failed packet alone using the Alamouti scheme, and P_2 will always decode the received retransmitted signal in Alamouti code regardless of cooperation of S_2. Based on the coding and modulation applied, the failed packet (D_f) would be translated into n baseband signals:

$$D(f): \{x_1, x_2, \dots, x_n\}.$$

Then the P_1 will transmit the following stream:

$$P_1: \{x_1, -x_2^*, x_3, -x_4^*, \dots, x_{n-1}, -x_n^*\},$$

and if S2 cooperates with P1, will transmit the following stream:

$$S_1: \{x_2, x_1^*, x_4, x_3^*, \dots, x_n, x_{n-1}^*\}.$$

which are the according to the Alamouti code. Then the receiver (P_2) will combine these signals according to (3.2) and will decode them according to the same scheme.

When the secondary network is active in the ARQ round, S_2 cooperates with P_1 in order to increase the power of the primary's signal to overcome that of interference introduced as a result of S_1's transmission. In other cases, when the secondary network is silent, P_1 retransmits the failed packet alone still in the Alamouti scheme where only one leg is active. Therefore, only the diversity gain is lost and the outcome is the same as if there was only one transmitter operating in normal (non-STBC) mode [84]. We could assume that P_1 retransmits the failed packet in Alamouti code only when S_2 is ready for cooperation, but this requires that P_1 monitors the secondary network's control signals, which is against the assumption of an oblivious primary network.

As the primary receiver node P_2 does not generate any NACK signals and negative declarations are based on lack of a positive (ACK) feedback, there would be no collision between CTS and feedback signals within the feedback time-slot. Otherwise, we would have to dedicate an extra time-slot for CTS alarms, or another frequency channel, which would increase the required bandwidth and degrade the overall throughput. The timing diagram shown in Fig. 3.4 depicts how the protocol works under different conditions of primary and secondary networks.

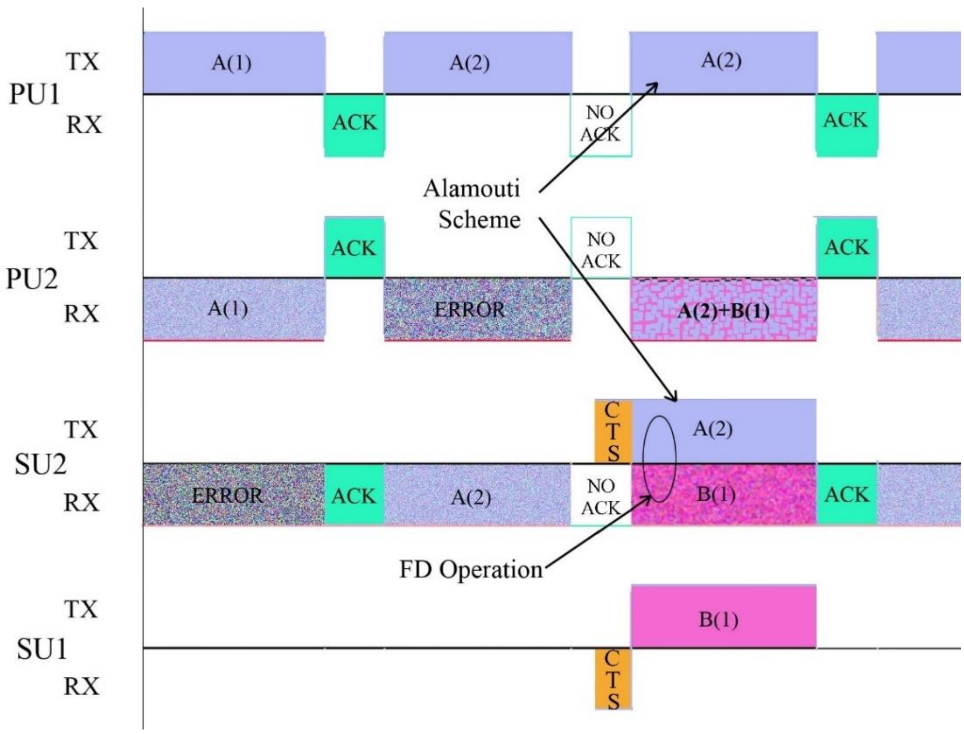


Fig. 3.4. Timing diagram of the system; primary & secondary networks are synchronized.

At the end of the ARQ round, P_2 tries to decode the received signal singly (Non-Hybrid ARQ) or together with the buffered signal of the original transmission (Hybrid ARQ). The primary network experiences different packet error rates and throughput for ARQ and HARQ modes, but it has no effect on the secondary's performance. On the other hand, the secondary's packet error rate and throughput are affected by the residual self-interference after cancellation (SIS factor β of S_2) which will all be analyzed in the following section.

The key features of this protocol are:

- The secondary pair communicates only during the ARQ rounds of the primary.
- The secondary receiver uses a perfect interference cancellation scheme (BKIC) which enables it to receive its cognitive message concurrently.
- In cases where the primary-on-secondary interference cannot be canceled (i.e. S_2 has not decoded the primary message correctly) it will not emit a CTS and the secondary pair remain silent. This is motivated by the idea that the operation of the secondary pair has costs (interference on the primary), therefore the secondary link should only be active when it can harvest good rates, otherwise it should stay silent.
- The primary network is oblivious to the presence of secondary users and will benefit from their cooperation where applicable without any coordination or synchronization.
- Exploiting the redundancy within ARQ rounds requires the combination of received signals during original and retransmission rounds which implies a Hybrid scheme. However, the use of the Alamouti code in our method which incorporates a diversity of order two shows promising results even in non-hybrid schemes which is a noticeable achievement.

3.4 Performance Analysis

3.4.1 Packet Error Rate

Although an analysis based on outage probability shows promising results in theory (as performed in [77]), however the practical results in a packet-based network are far from it. In practice we are faced with finite number of modulations and coding schemes and large-sized packets that are far from ideal assumptions made in outage-based analyses [113]. Moreover, in an ARQ-based system, a secondary signal may cause a packet error in the primary network which, if compensated by ARQ retransmissions, will not introduce an outage, but will degrade the throughput of the system. Therefore, the outage probability alone, is not a proper measure for the performance of a network. On the other hand, MAC layer Packet Error Rate (PER) is a more practical and applicable metric for design paradigms, and in this section we will consider this metric in evaluating performance of the proposed scheme.

In this section we use the approximate value for PER given in (3.5) which is a function of the system's signal-to-noise ratio. In a Rayleigh fading channel, signal-to-noise ratio of the received signal at the receiver γ follows an exponential distribution with the following probability density function:

$$f_\gamma(\gamma) = \frac{1}{\bar{\gamma}} \exp\left(\frac{-\gamma}{\bar{\gamma}}\right), \tag{3.6}$$

where $\bar{\gamma}$ is the average received signal-to-noise ratio.

Non-Hybrid ARQ Mode:

For a truncated ARQ system with one retransmission attempt (as assumed in our proposed scheme), the overall packet error rate (denoted by $\mathcal{E}$) is the product of packet error rate of original and repeated packets:

$$\mathcal{E}(\gamma_I, \gamma_R) = PER(\gamma_I)PER(\gamma_R), \tag{3.7}$$

where γ_I and γ_R are the signal-to-noise ratios of the initial and repeated packets, respectively. For traditional ARQ (without cooperation) in our system model, both the initial and repeated packets are transmitted by P_1, and for a slow fading channel we can assume $\gamma_I = \gamma_R = \gamma_{P_1,P_2}$ where γ_{P_1,P_2} is the instantaneous signal-to-noise ratio at the receiver of P_2 corresponding to the signal transmitted by P_1. Therefore, equation (3.7) reduces to:

$$\mathcal{E}_T(\gamma_{\mathrm{P}_1,\mathrm{P}_2}) = PER^2(\gamma_{\mathrm{P}_1,\mathrm{P}_2}), \tag{3.8}$$

$\mathcal{E}_T$ denotes traditional (non-cooperative) ARQ protocol. The average value of $\mathcal{E}_T$ for an ergodic channel would be:

$$\begin{aligned}\bar{\mathcal{E}}_T = \mathbb{E}\left[PER^2(\gamma_{\mathrm{P}_1,\mathrm{P}_2})\right] &= \int_0^\infty PER^2(\gamma_{\mathrm{P}_1,\mathrm{P}_2})\, f_{\gamma_{\mathrm{P}_1,\mathrm{P}_2}}(\gamma_{\mathrm{P}_1,\mathrm{P}_2})\, d\gamma_{\mathrm{P}_1,\mathrm{P}_2} \\ &= 1 - \frac{2\bar{\gamma}_{\mathrm{P}_1,\mathrm{P}_2} g}{1 + 2\bar{\gamma}_{\mathrm{P}_1,\mathrm{P}_2}} \exp\left(\frac{-\gamma_t}{\bar{\gamma}_{\mathrm{P}_1,\mathrm{P}_2}}\right),\end{aligned} \tag{3.9}$$

where $f_{\gamma_{\mathrm{P}_1,\mathrm{P}_2}}(\gamma_{\mathrm{P}_1,\mathrm{P}_2})$ represents the pdf of $\gamma_{\mathrm{P}_1,\mathrm{P}_2}$. It should be noted that the packet error rate of the original and repeated packets in cases where S_1 does not cooperate in the ARQ process have the same value, although the coding in the retransmission round (by P_1 alone) is STBC (Alamouti) and differs from the original round. As explained earlier, in an Alamouti coded system, when a soft failure occurs (one of the transmitting antennas fails to transmit), only the diversity gain is lost and the signal will be detected as it had been transmitted with just one antenna without STBC (Aalamouti) coding with similar performance [84].

According to the proposed cooperative protocol, S_2 will cooperate in retransmission of the initial packet if it has decoded the initial transmission correctly. Otherwise P_1 will repeat the packet alone, i.e. without cooperation with S_2. In the former case, the P2 receiver will benefit from diversity gain of Alamouti transmission resulting in a lower PER. In the latter case (no cooperation from S_2), there is no diversity gain, so the PER would be higher. The average probability that S2 receives the packets without error (average packet success rate $\overline{PSR}$) is:

$$\overline{PSR}(\gamma_{P_1,S_2}) = 1 - \overline{PER}(\gamma_{P_1,S_2}) = 1 - \int_0^\infty PER(\gamma_{P_1,S_2})\, f_{\gamma_{P_1,S_2}}(\gamma_{P_1,S_2})\, d\gamma_{P_1,S_2}$$
$$= \frac{\bar{\gamma}_{P_1,S_2} g}{1 + \bar{\gamma}_{P_1,S_2} g} \exp\left(\frac{-\gamma_t}{\bar{\gamma}_{P_1,S_2}}\right). \tag{3.10}$$

From now on we denote γ_{P_1,P_2} by γ_A, γ_{S_2,P_2} by γ_B, γ_{S_1,P_2} by γ_C, $\gamma_{S_1,S_2}$by γ_D and γ_{P_1,S_2} by γ_E for brevity, as shown in Fig 3.5.

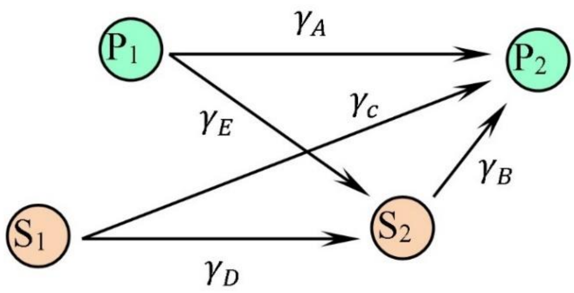


Fig. 3.5. Signal-to-noise ratios of the channels of the network.

In cooperative mode, P1 and S2 repeat the packet using the Alamouti code. Therefore, the instantaneous SNR at the receiver of P2 for the repeated packet will be $\gamma_R = \frac{\gamma_A + \gamma_B}{1+\gamma_C}$ in ARQ mode. Hence, the probability that a packet fails in both initial and cooperative repeat rounds in one round of ARQ is:

$$\mathcal{E}_C(\gamma_A, \gamma_B, \gamma_C) = PER(\gamma_A) PER\left(\frac{\gamma_A + \gamma_B}{1 + \gamma_C}\right), \tag{3.11}$$

and its average value would be:

$$\bar{\mathcal{E}}_C = \mathbb{E}[\mathcal{E}_C(\gamma_A, \gamma_B, \gamma_C)]$$
$$= \iiint_0^\infty PER(\gamma_A) PER\left(\frac{\gamma_A + \gamma_B}{1 + \gamma_C}\right) f_{\gamma_A}(\gamma_A) f_{\gamma_B}(\gamma_B) f_{\gamma_C}(\gamma_C) d\gamma_A . d\gamma_B . d\gamma_C. \tag{3.12}$$

$f_{\gamma_A}(\gamma_A)$, $f_{\gamma_B}(\gamma_B)$, and $f_{\gamma_C}(\gamma_C)$ are pdfs of γ_A, γ_B, and γ_C respectively and $\bar{\mathcal{E}}_C$ denotes the average packet error rate in cooperative ARQ. To the best of our knowledge, the above integral has no exact closed form. But we can simplify the above integral by a valid assumption. As we have assumed all codewords are drawn from a Guassian codebook, we may consider the interference γ_C as a noise (with zero mean and variance $\bar{\gamma}_C$), and reduce the dimensions of the above integral to two:

$$\bar{\mathcal{E}}_C = \iint_0^\infty PER(\gamma_A) PER(\frac{\gamma_A + \gamma_B}{1 + \bar{\gamma}_C}) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma) d\gamma_A . d\gamma_B . \tag{3.13}$$

Now we can calculate the above integral over proper boundaries. This integration should be carried out over 5 different boundaries where the values of $PER(\gamma_A)$ and $PER(\frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C})$ switch from 1 to an exponential function according to (3.5). The respective boundaries are as follows:

$$\begin{aligned}
&\int_0^{\gamma_t} \int_0^{\gamma_t(1+\bar{\gamma}_C)-\gamma_A} \Phi(\gamma_A, \gamma_B) d\gamma_A . d\gamma_B + \int_0^{\gamma_t} \int_{\gamma_t(1+\bar{\gamma}_C)-\gamma_A}^{\infty} \Phi(\gamma_A, \gamma_B) d\gamma_A . d\gamma_B \\
&+ \int_{\gamma_t}^{\gamma_t(1+\bar{\gamma}_C)} \int_0^{\gamma_t(1+\bar{\gamma}_C)-\gamma_A} \Phi(\gamma_A, \gamma_B) d\gamma_A . d\gamma_B + \int_{\gamma_t}^{\gamma_t(1+\bar{\gamma}_C)} \int_{\gamma_t(1+\bar{\gamma}_C)-\gamma_A}^{\infty} \Phi(\gamma_A, \gamma_B) d\gamma_A . d\gamma_B \\
&+ \int_{\gamma_t(1+\bar{\gamma}_C)}^{\infty} \int_0^{\infty} \Phi(\gamma_A, \gamma_B) d\gamma_A . d\gamma_B
\end{aligned} \tag{3.14}$$

where $\Phi(\gamma_A, \gamma_B) = PER(\gamma_A) PER(\frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C}) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$, and its value within the boundaries of any of the above five integrals would be as in Table 3.5.

TABLE 3.5. The values of the integrand in (3.14) within respective boundaries.

First integral	$\Phi(\gamma_A, \gamma_B) = f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Second integral	$\Phi(\gamma_A, \gamma_B) = \alpha \exp\left(-g \frac{\gamma_A + \gamma_B}{1 + \bar{\gamma}_C}\right) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Third integral	$\Phi(\gamma_A, \gamma_B) = \alpha \exp(-g\gamma_A) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Fourth Integral	$\Phi(\gamma_A, \gamma_B) = \alpha^2 \exp(-g\gamma_A) \exp\left(-g \frac{\gamma_A + \gamma_B}{1 + \bar{\gamma}_C}\right) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Fifth integral	$\Phi(\gamma_A, \gamma_B) = \alpha^2 \exp(-g\gamma_A) \exp\left(-g \frac{\gamma_A + \gamma_B}{1 + \bar{\gamma}_C}\right) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$

The result of integration of (3.13) has been shown in (3.15).

$$
\begin{aligned}
\bar{\mathcal{E}}_C = {} & 1 - exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) - \frac{\bar{\gamma}_B^2 g}{(\bar{\gamma}_B g + \bar{\gamma}_C + 1)(\bar{\gamma}_A - \bar{\gamma}_B)}\left(exp\left(\frac{-\gamma_t\bar{\gamma}_C}{\bar{\gamma}_B} - \frac{\gamma_t}{\bar{\gamma}_A}\right) - exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_B}\right)\right) \\
& + \frac{\bar{\gamma}_B}{\bar{\gamma}_A - \bar{\gamma}_B - \bar{\gamma}_A\bar{\gamma}_B g}\left(exp\left(\frac{-\gamma_t\bar{\gamma}_C}{\bar{\gamma}_B} - \frac{\gamma_t}{\bar{\gamma}_A}\right) - \alpha^{-\bar{\gamma}_C} exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A}\right)\right) + \frac{1}{(1+\bar{\gamma}_A g)} exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) \\
& - \frac{\alpha^{-\bar{\gamma}_C}}{(1+\bar{\gamma}_A g)} exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A}\right) + \frac{\alpha^{-\bar{\gamma}_C}(1+\bar{\gamma}_C)^2}{(1+\bar{\gamma}_C+\bar{\gamma}_B g)\big((1+\bar{\gamma}_C)(1+\bar{\gamma}_A g)+\bar{\gamma}_A g\big)} exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A}\right) \\
& + \frac{\bar{\gamma}_B(1+\bar{\gamma}_C)}{(1+\bar{\gamma}_C+\bar{\gamma}_B g)(\bar{\gamma}_A - \bar{\gamma}_B - \bar{\gamma}_A\bar{\gamma}_B g)}\left(\alpha^{-\bar{\gamma}_C} exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A}\right) - exp\left(\frac{-\gamma_t\bar{\gamma}_C}{\bar{\gamma}_B} - \frac{\gamma_t}{\bar{\gamma}_A}\right)\right)
\end{aligned}
\tag{3.15}
$$

Re ordering or factorization of the above equation does not add to its simplicity. But we will consider some special cases to simplify it more and get a better insight. When the signal-to-noise ratio of the interference signal is more than 1 dB (i.e. $\bar{\gamma}_C > 1$) we can ignore the terms containing $\alpha^{-\bar{\gamma}_C}$ (according to Tables 3.2 and 3.3 α is a two-digit number). Thus, with this assumption we can rewrite (3.15) as follows:

$$
\begin{aligned}
\bar{\mathcal{E}}_C = {} & 1 - \frac{\bar{\gamma}_A g}{(1+\bar{\gamma}_A g)} exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) + \frac{\bar{\gamma}_B^2 g}{(\bar{\gamma}_B g + \bar{\gamma}_C + 1)(\bar{\gamma}_A - \bar{\gamma}_B)} exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A}\right) \\
& + \frac{\bar{\gamma}_A\bar{\gamma}_B^2 g^2}{(\bar{\gamma}_B g + \bar{\gamma}_C + 1)(\bar{\gamma}_A - \bar{\gamma}_B)(\bar{\gamma}_A - \bar{\gamma}_B - \bar{\gamma}_A\bar{\gamma}_B g)} exp\left(\frac{-\gamma_t\bar{\gamma}_C}{\bar{\gamma}_B} - \frac{\gamma_t}{\bar{\gamma}_A}\right)
\end{aligned}
\tag{3.16}
$$

and if we further assume that the SNRs of the primary and secondary transmissions are the same at P_2's receiver (i.e. $\bar{\gamma}_A = \bar{\gamma}_B$), it would be

$$
\bar{\mathcal{E}}_C = 1 - \frac{\bar{\gamma}_A g}{(1+\bar{\gamma}_A g)} exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) - \frac{1+\gamma_t g}{(\bar{\gamma}_A g + \bar{\gamma}_C + 1)} exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A}\right). \tag{3.17}
$$

If during the ARQ round S_1 does not transmit its message despite a CTS, and S_2 keeps cooperating (i.e. $\bar{\gamma}_A = \bar{\gamma}_B$ and $\bar{\gamma}_C = 0$), (3.17) will further simplify to the more compact form which is provided for comparison purposes:

$$
\bar{\mathcal{E}}_C = 1 - exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) - \frac{\gamma_t g}{(1+\bar{\gamma}_A g)} exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) + \frac{1}{(1+\bar{\gamma}_A g)(1+2\bar{\gamma}_A g)} exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right). \tag{3.18}
$$

Hybrid ARQ (HARQ) Mode

In this section, we consider Hybrid ARQ with incremental redundancy which means the same copy of the original packet is retransmitted in the ARQ round and the receiver combines the signals corresponding to the failed packet (buffered in P_2), and the retransmitted packet in the ARQ round to increase the mutual information.

Therefore, the effective SNR in HARQ decoding can be assumed to be $\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C}$. Then we derive the corresponding packet error rate as follows:

$$\bar{\mathcal{E}}_{CH} = \iint_0^\infty PER(\gamma_A)PER(\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C})\, f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma) d\gamma_A . d\gamma_B . \tag{3.19}$$

Here the integration should be carried out over four different boundaries where according to (3.5) the values of $PER(\gamma_A)$ and $PER(\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C})$ switch from 1 to an exponential function. The respective boundaries are as follows:

$$\int_0^{\frac{\gamma_t(1+\bar{\gamma}_C)}{(2+\bar{\gamma}_C)}} \int_0^{\gamma_t(1+\bar{\gamma}_C)-\gamma_A(2+\bar{\gamma}_C)} \Psi(\gamma_A,\gamma_B) d\gamma_A . d\gamma_B$$

$$+ \int_0^{\frac{\gamma_t(1+\bar{\gamma}_C)}{(2+\bar{\gamma}_C)}} \int_{\gamma_t(1+\bar{\gamma}_C)-\gamma_A(2+\bar{\gamma}_C)}^{\infty} \Psi(\gamma_A,\gamma_B) d\gamma_A . d\gamma_B$$

$$+ \int_{\frac{\gamma_t(1+\bar{\gamma}_C)}{(2+\bar{\gamma}_C)}}^{\gamma_t} \int_0^\infty \Psi(\gamma_A,\gamma_B) d\gamma_A . d\gamma_B + \int_{\gamma_t}^\infty \int_0^\infty \Psi(\gamma_A,\gamma_B) d\gamma_A . d\gamma_B \tag{3.20}$$

where $\Psi(\gamma_A,\gamma_B) = PER(\gamma_A)PER(\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C}) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$, and its value for each of the above four integrals, considering the respective boundaries, would be as according to Table 3.6.

TABLE 3.6. The value of the integrand in (3.20) within respective boundaries.

First integral	$\Psi(\gamma_A,\gamma_B) = f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Second integral	$\Psi(\gamma_A,\gamma_B) = \alpha \exp\left(-g(\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C})\right) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Third integral	$\Psi(\gamma_A,\gamma_B) = \alpha \exp\left(-g(\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C})\right) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$
Fourth integral	$\Psi(\gamma_A,\gamma_B) = \alpha^2 \exp(-g\gamma_A) \exp\left(-g(\gamma_A + \frac{\gamma_A+\gamma_B}{1+\bar{\gamma}_C})\right) f_{\gamma_A}(\gamma) f_{\gamma_B}(\gamma)$

The result of integration of (3.20) is shown in (3.21):

$$
\begin{aligned}
\bar{\mathcal{E}}_{CH} = 1 - \exp\left(\frac{-(1+\bar{\gamma}_C)\gamma_t}{\bar{\gamma}_A(2+\bar{\gamma}_C)}\right) \\
- \frac{\bar{\gamma}_B^2 g(2+\bar{\gamma}_C)}{(1+\bar{\gamma}_C+\bar{\gamma}_B\bar{\gamma}_C g+2\bar{\gamma}_B g)\big((2+\bar{\gamma}_C)\bar{\gamma}_A-\bar{\gamma}_B\big)}\left(\exp\left(\frac{-(1+\bar{\gamma}_C)\gamma_t}{\bar{\gamma}_A(2+\bar{\gamma}_C)}\right)-\exp\left(\frac{-(1+\bar{\gamma}_C)\gamma_t}{\bar{\gamma}_B}\right)\right) \\
- \frac{(1+\bar{\gamma}_C)^2}{(1+\bar{\gamma}_C+\bar{\gamma}_A\bar{\gamma}_C g+2\bar{\gamma}_A g)(1+\bar{\gamma}_C+\bar{\gamma}_B\bar{\gamma}_C g+2\bar{\gamma}_B g)}\left(\frac{1}{L}\exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right)-\exp\left(\frac{-(1+\bar{\gamma}_C)\gamma_t}{\bar{\gamma}_A(2+\bar{\gamma}_C)}\right)\right) \\
+ \frac{(1+\bar{\gamma}_C)}{L(1+\bar{\gamma}_C+\bar{\gamma}_B\bar{\gamma}_C g+2\bar{\gamma}_B g)(1+\bar{\gamma}_C+\overline{2\gamma_A}\bar{\gamma}_C g+3\bar{\gamma}_A g)}\exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right),
\end{aligned}
\tag{3.21}
$$

where $L = \sqrt[(1+\bar{\gamma}_C)]{\alpha}$. In calculating the above integrals, we have used the switching threshold relation $\alpha \exp(-g\gamma_t) = 1$ whenever applicable. In order to have a more meaningful form of the above equation we simplify it for special cases. When $\bar{\gamma}_A = \bar{\gamma}_B$ and for high values of $\bar{\gamma}_C$ we can assume $2+\bar{\gamma}_C \approx 1+\bar{\gamma}_C$ and rewrite (3.21) as follows:

$$
\bar{\mathcal{E}}_{CH} = 1 - M\exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right) + N\exp\left(\frac{-(1+\bar{\gamma}_C)\gamma_t}{\bar{\gamma}_A}\right) \tag{3.22}
$$

where M and N are defined as follows:

$$
\begin{gathered}
M = \frac{\bar{\gamma}_A g(1+2\bar{\gamma}_A g+3\bar{\gamma}_C+2\bar{\gamma}_C\bar{\gamma}_A g)}{\bar{\gamma}_C(1+\bar{\gamma}_A g)(1+2\bar{\gamma}_A g)} \\
N = \frac{\bar{\gamma}_A g}{(1+\bar{\gamma}_A g)}
\end{gathered}
\tag{3.23}
$$

Having found the average error rate of a packet after one original and one repeat round, we can calculate the overall packet error rate of the system in both ARQ and HARQ modes.

The overall packet error rate of the cooperative protocol in the non-Hybrid ARQ mode will be as follows:

$$
P_P^e = \big(1-\overline{PSR}(\gamma_E)\big)\bar{\mathcal{E}}_T + \overline{PSR}(\gamma_E)\bar{\mathcal{E}}_C, \tag{3.24}
$$

where P_P^e is the average PER of the primary network. For the Hybrid ARQ mode, we need to replace $\bar{\mathcal{E}}_C$ by $\bar{\mathcal{E}}_{CH}$ in (3.24):

$$
P_P^e = \big(1-\overline{PSR}(\gamma_E)\big)\bar{\mathcal{E}}_T + \overline{PSR}(\gamma_E)\bar{\mathcal{E}}_{CH}, \tag{3.25}
$$

For comparison purposes, we derive the average packet error rate of non-cooperative half duplex scheme (proposed in [77]) as follows:

$$\bar{\mathcal{E}}_{CH} = \int_0^{\infty} PER(\gamma_A) PER\left(\gamma_A + \frac{\gamma_A}{1+\bar{\gamma}_C}\right) f_{\gamma_A}(\gamma) d\gamma_A =$$

$$\int_0^{\frac{\gamma_t(1+\bar{\gamma}_C)}{(2+\bar{\gamma}_C)}} \Lambda(\gamma_A) d\gamma_A + \int_{\frac{\gamma_t(1+\bar{\gamma}_C)}{(2+\bar{\gamma}_C)}}^{\gamma_t} \Lambda(\gamma_A) d\gamma_A + \int_{\gamma_t}^{\infty} \Lambda(\gamma_A) d\gamma_A =$$

$$1 - \frac{\bar{\gamma}_A g(2+\bar{\gamma}_C)}{(1+\bar{\gamma}_A g)(1+\bar{\gamma}_C) + \bar{\gamma}_A g} \exp\left(\frac{-\gamma_t(1+\bar{\gamma}_C)}{\bar{\gamma}_A(2+\bar{\gamma}_C)}\right) \tag{3.26}$$

$$- \frac{\bar{\gamma}_A g(1+\bar{\gamma}_C)^2}{\big((1+\bar{\gamma}_A g)(1+\bar{\gamma}_C) + \bar{\gamma}_A g\big)\big((1+2\bar{\gamma}_A g)(1+\bar{\gamma}_C) + \bar{\gamma}_A g\big)} \times \exp\left(\frac{-\bar{\gamma}_t g}{(1+\bar{\gamma}_C)}\right) \exp\left(\frac{-\gamma_t}{\bar{\gamma}_A}\right)$$

where $\Lambda(\gamma_A) = PER(\gamma_A) PER\left(\gamma_A + \frac{\gamma_A}{1+\bar{\gamma}_C}\right) f_{\gamma_A}(\gamma)$.

For the secondary network we have only one round of transmission. Therefore, the MAC layer packet error rate for secondary transmission (P_S^e) can be derived by calculating the average value of (3.5) where the instantaneous SNR is set to $\frac{\gamma_D}{1+\beta\gamma_B}$:

$$P_S^e = 1 - \frac{g\bar{\gamma}_D}{1+\beta\bar{\gamma}_B + g\bar{\gamma}_D} \exp\left(\frac{-(1+\beta\bar{\gamma}_B)\gamma_t}{\bar{\gamma}_D}\right) \tag{3.27}$$

3.4.2 Throughput

One of the critical metrics in any communication network is the throughput; the amount of information transferred within a specific time. Having found the overall packet error rate of the primary and secondary networks in the previous section, derivation of the respective throughputs would be straightforward. For the proposed method, the primary's throughput would be:

$$R_p = R_{p_0}(1 - P_P^e) \tag{3.28}$$

In the secondary network, transmission will happen when a primary packet has failed, and at the same time it has been received correctly by SU_2. Therefore, secondary throughput would be:

$$R_s = R_{s_0} \overline{PER}(\gamma_A) \overline{PSR}(\gamma_E)(1 - P_S^e) \tag{3.29}$$

3.5 Simulation Results

In this section, we present some simulation results to demonstrate the performance of the proposed schemes. For simulation purposes we have considered BPSK modulation and have used the parameters provided in Table 3.2 and Table 3.3 for a packet length of 1080 bits, as in HIPERLAN and 802.11a standards. Similar performance for other modulation schemes would be expected. We have considered both ARQ and HARQ schemes and compared the results with non-cooperative half duplex scheme presented in [77]. We have assumed that the channels between any two nodes (primary or secondary) are all Rayleigh flat fading with similar conditions and approximately similar distance for simulation purposes.

Fig. 3.6 shows the overall PER for the primary network of the proposed scheme in ARQ and HARQ modes compared to traditional non-cooperative ARQ scheme and non-cooperative half-duplex HARQ scheme offered in [77] when primary and secondary transmitters use the same power levels ($P_P = P_s$ which is equivalent to $\bar{\gamma}_A = \bar{\gamma}_B$ in our simulation scenario). As can be seen our proposed scheme for ARQ and

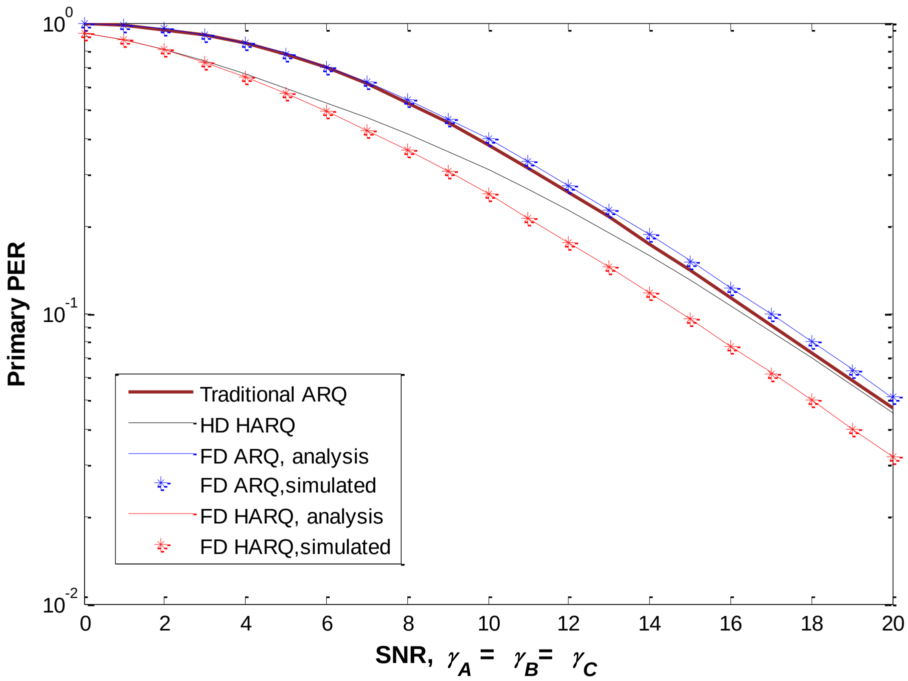


Fig. 3.6. Primary PER for ARQ and HARQ modes for equal primary and secondary powers, un-coded BPSK modulation.

HARQ modes achieves slightly higher PER compared to the HD traditional non-cooperative ARQ scheme, which could be anticipated. When powers are equal in the primary and secondary networks, the average SNR of the received signal in ARQ round is a small amount (app. 2dB) which will not contribute significantly to the overall PER enhancement. However, our HARQ scheme shows considerable enhancement over the non-cooperative half duplex HARQ method investigated in other studies. As we see, the non-cooperative half duplex HARQ has degraded the PER performance of the primary network, whereas our proposed cooperative FD scheme improves it in the worst case when $\bar{\gamma}_A = \bar{\gamma}_B = \bar{\gamma}_c$.

Fig. 3.7 shows the same metric for the case when the secondary power is half of primary power ($\bar{\gamma}_A = \bar{\gamma}_B = 2\bar{\gamma}_c$) and consequently less interference from a secondary user is imposed on the primary's signal. As it is expected, the primary's performance is considerably enhanced, and for SNR ratios above 12 dB, we observe approximately 2dB savings for Hybrid ARQ mode compared to the half duplex HARQ scheme proposed in [77]. We see a perfect match between analytical and simulation results in Fig. 3.6 and Fig. 3.7 which verifies the validity of our analyses. This improvement in

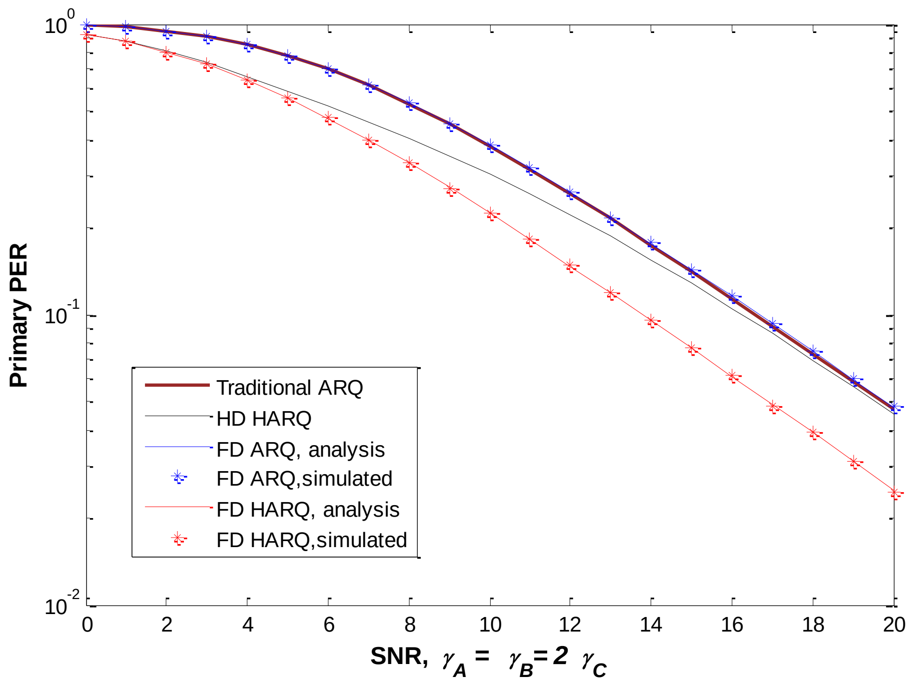


Fig. 3.7. Primary PER for ARQ and HARQ modes for unequal primary and secondary powers, un-coded BPSK modulation.

packet error rate of the primary network would be more if we decrease the secondary user's transmission power P_s further.

When using coded modulations, the system is more robust against errors, and better results are anticipated for our protocol compared to the traditional ARQ protocol and half duplex HARQ scheme. In Fig. 3.8. we have shown PER of the primary network for a convolutionally coded BPSK signal with coding rate ½, where primary and secondary powers are equal ($\bar{\gamma}_A = \bar{\gamma}_B = \bar{\gamma}_C$). We observe an enhanced PER behavior for the coded modulation scheme compared to the uncoded modulation. In addition, we observe that even in ARQ mode, our scheme outperforms the traditional ARQ and half duple HARQ schemes considerably. For a target PER, we have app. 6dB savings in transmitters power for our scheme compared to half duplex ARQ, which is a significant achievement. Fig. 3.9 depicts the same metric, when the primary and secondary powers are not equal ($\bar{\gamma}_A = \bar{\gamma}_B = 2\bar{\gamma}_C$). And as expected, the performance is much more enhanced.

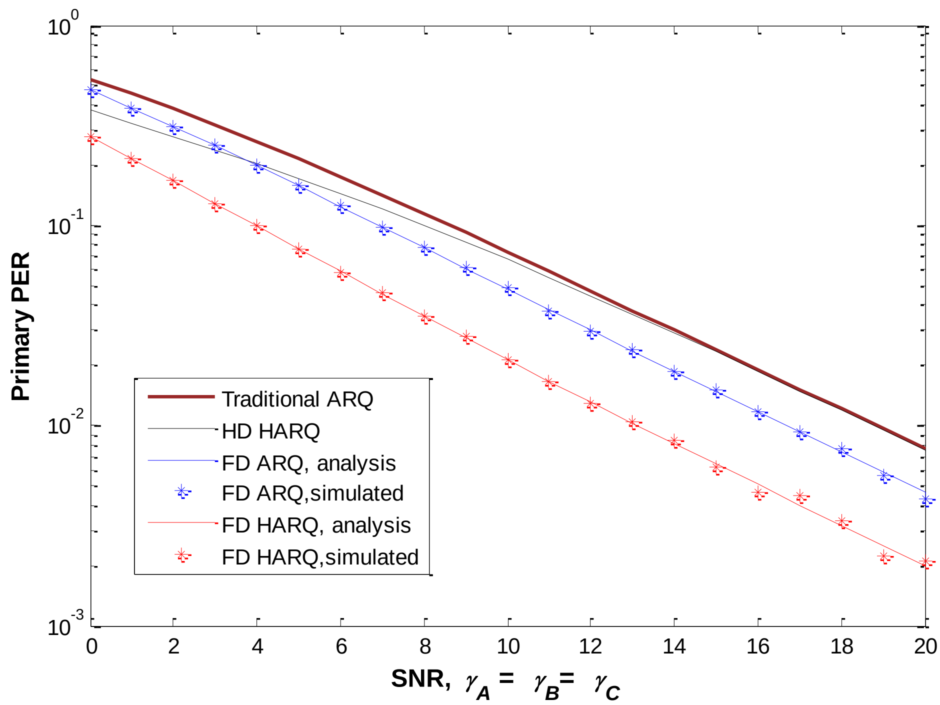


Fig. 3.8. Primary PER for ARQ and HARQ modes for equal primary and secondary powers, convolutionally coded BPSK modulation.

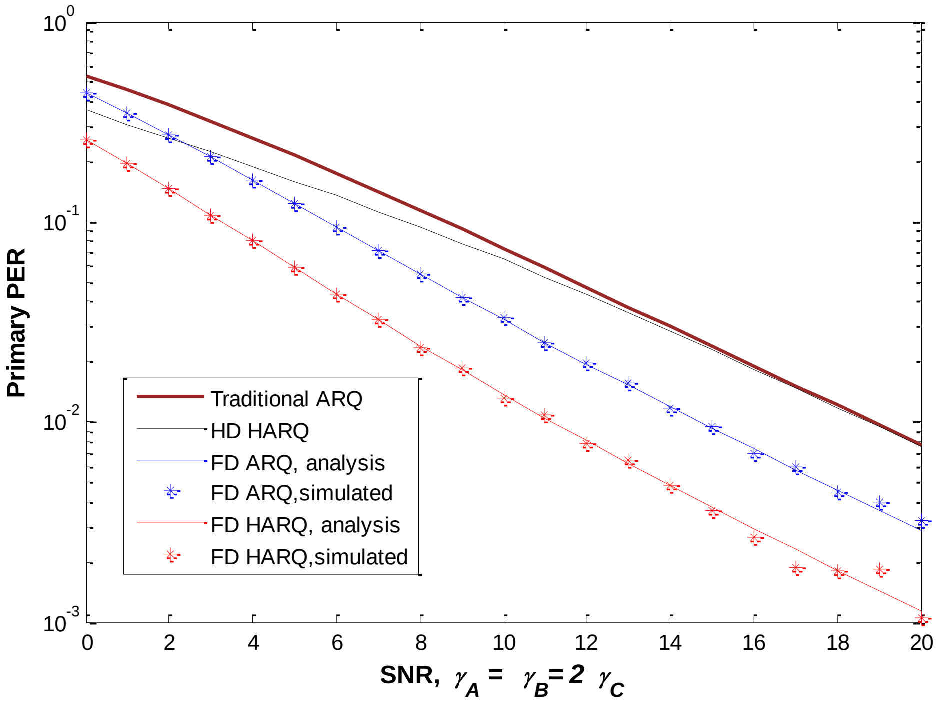


Fig. 3.9. Primary PER for ARQ and HARQ modes for unequal primary and secondary powers, convolutionally coded BPSK modulation.

In Fig. 3.10 we have depicted the normalized throughput of the primary network with an uncoded BPSK modulation in the proposed scheme for ARQ and HARQ modes, compared to that in the traditional non-cooperative ARQ as well as half duplex HARQ scheme presented in [77] when $P_P = P_S$. Fig. 3.11 shows the same metric when $P_P = 2P_S$. From both figures, it is obvious that FD cooperative HARQ enhances the throughput compared to traditional non-cooperative half duplex HARQ (proposed in [77]), although the improvement is much more when $P_P = 2P_S$. For equal primary and secondary power case (Fig. 3.10.) the FD ARQ scheme is not better than HD HARQ, or even traditional non-cooperative ARQ mode. The 2dB SNR of the received signal at P_2 is not enough to compensate for the interference caused by S_1 transmission, when the primary and secondary powers are the same. When transmission power of S_1 is less than that of P_1, the overall SNR at P_2 increases and the PER performance of primary network improves. When $P_P = 2P_S$ the performance of FD ARQ is nearly the same as traditional non-cooperative ARQ scheme as shown in Fig. 3.11.

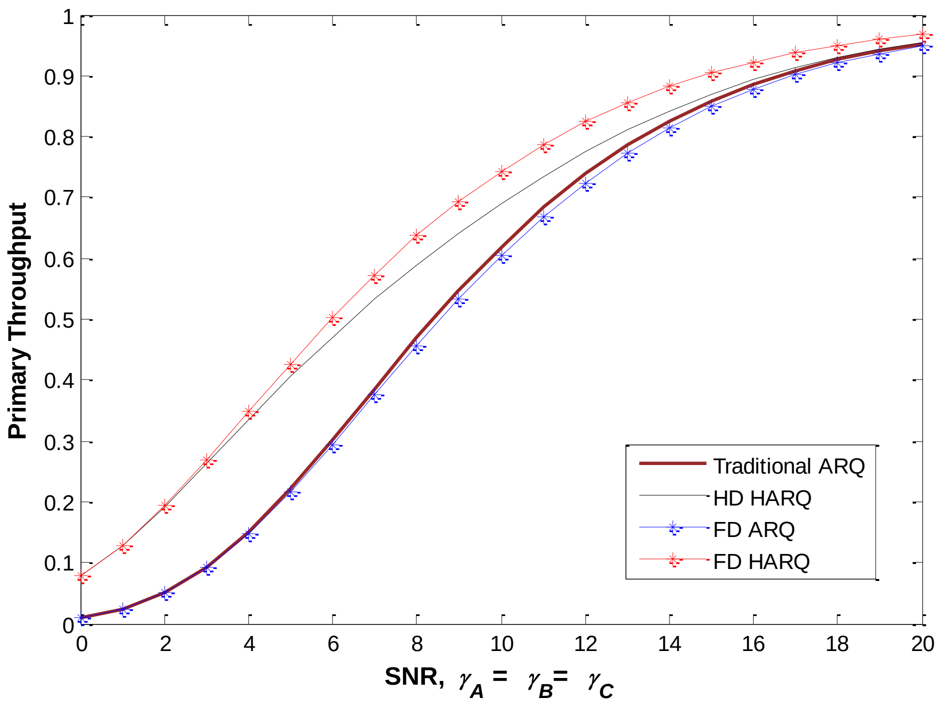


Fig. 3.10. Primary throughput for equal primary and secondary powers ($\bar{\gamma}_A = \bar{\gamma}_B = \bar{\gamma}_c$), uncoded BPSK modulation.

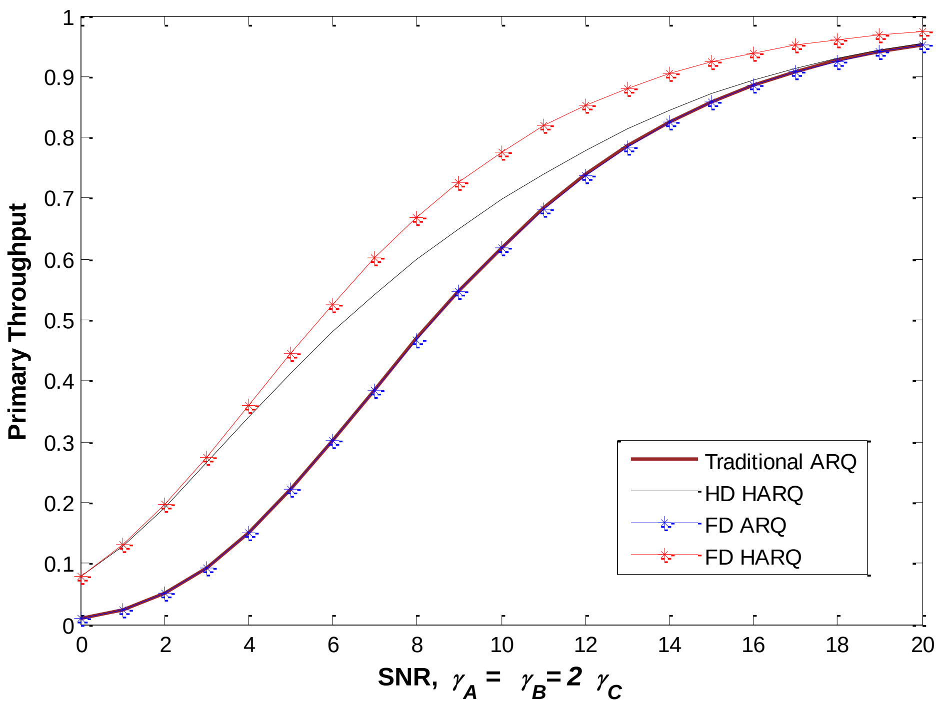


Fig. 3.11. Primary throughput for unequal primary and secondary powers ($\bar{\gamma}_A = \bar{\gamma}_B = 2\bar{\gamma}_c$), uncoded BPSK modulation.

Figures 3.12 and 3.13 show the primary's PER for convolutionally code BPSK modulation for equal and unequal primary and secondary powers respectively.

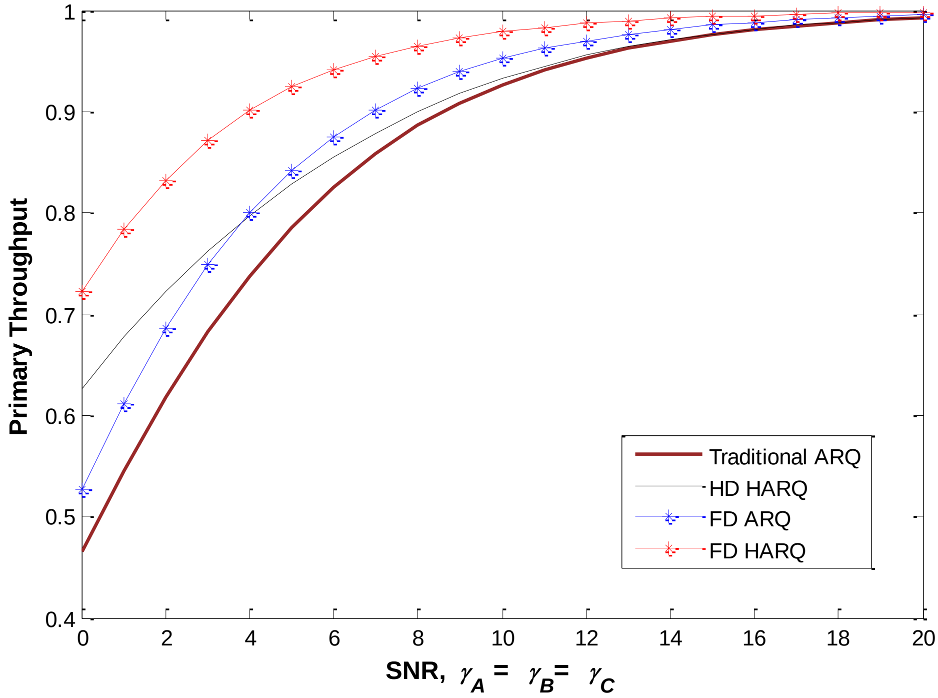


Fig. 3.12. Primary throughput for equal primary and secondary powers ($\bar{\gamma}_A = \bar{\gamma}_B = \bar{\gamma}_c$), convolutionally coded BPSK modulation.

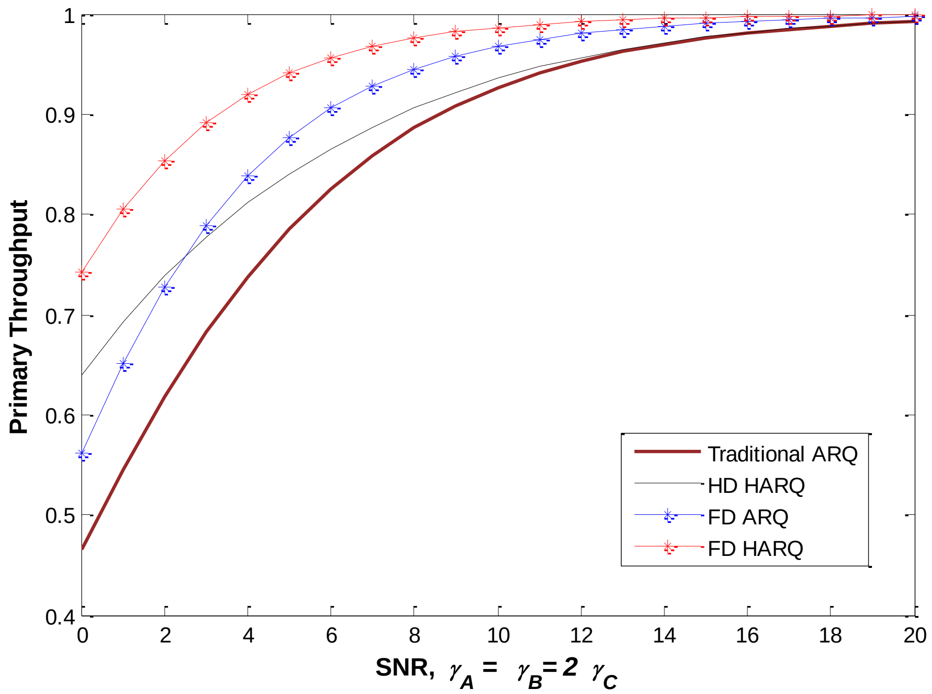


Fig. 3.13. Primary throughput for unequal primary and secondary powers ($\bar{\gamma}_A = \bar{\gamma}_B = 2\bar{\gamma}_c$), convolutionally coded BPSK modulation.

When coded modulation is implemented, the system is more reliable, and less prone to errors, which results in the better performance of the proposed scheme. Here the FD ARQ scheme shows a superior performance to that of traditional ARQ for all SNRs and to HD ARQ scheme at SNRs more than 5dB when $P_P = P_S$ (Fig. 3.12), and 3dB when $P_P = 2P_S$ (Fig. 3.13). For FD HARQ mode, we observe a 4dB savings in transmission power at low SNRs, which is a significant achievement.

Apart from the primary network's performance metrics, the secondary networks PER and throughput are of interest which are attended to in the following.

As we have assumed perfect KIC capability for S_2, the power and interference of the primary network on S_2 receiver has no direct effect on packet error rate of the secondary network. However, since we have assumed that power of S_2 is the same as P_1 transmitter (i.e. P_P) in order to achieve a diversity of order two through Alamouti scheme transmission, the primary power indirectly affects the average PER of the secondary network through residual self-interference pertaining to full duplex operation of S_2. In other words, when the power of P_1 changes, S_2 follows it and change its transmission power accordingly.

In Fig. 3.14 and Fig. 3.15 we have depicted the average packet error rate of the secondary network versus SNR of the secondary network at the receiver of S_2, for different SIS capability factors when $P_P = P_S$ and when $P_P = 2P_S$ respectively. When SIS is perfect (i.e. $\beta = 0$), the power of the primary network has no effect on the secondary's PER, as observed from these figures. But, when SIS is not perfect, as the transmission power of S_1 increases, the effect of residual self-interference is noticeable and is dominant for high values of the secondary SNR. The increase in PER against SIS factor is not linear which is observed in the figure. As it is seen the simulation results follow the analytical graph closely which verifies validity of our analyses.

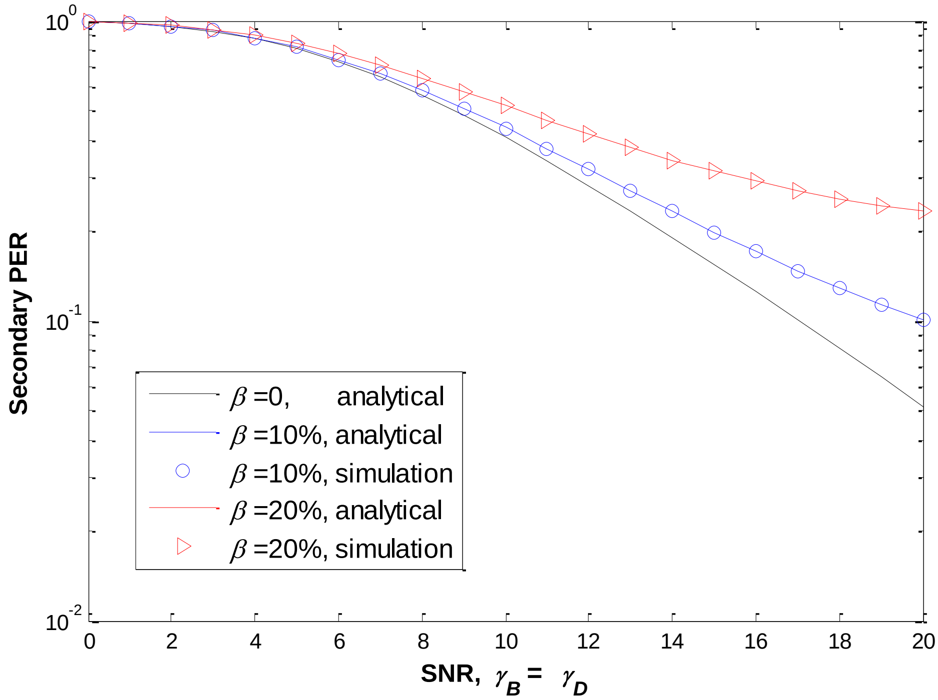


Fig. 3.14. Secondary PER vs. secondary SNR; equal primary and secondary powers ($\bar{\gamma}_B = \bar{\gamma}_D$).

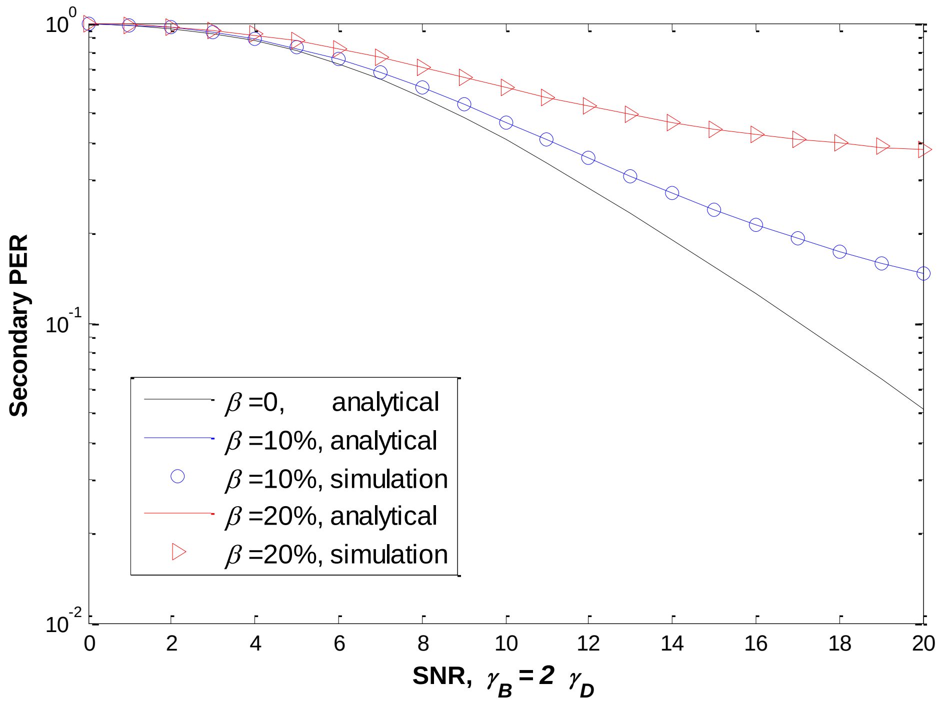


Fig. 3.15. Secondary PER vs. secondary SNR, unequal primary and secondary powers ($\bar{\gamma}_B = \bar{\gamma}_D$).

Fig. 3.16 depicts the changes in secondary PER vs. SIS factor (β) for equal and unequal primary and secondary power scenarios. As expected, the secondary's PER will increase when the residual self-interference increases, and as mentioned in previous section, it is not linear.

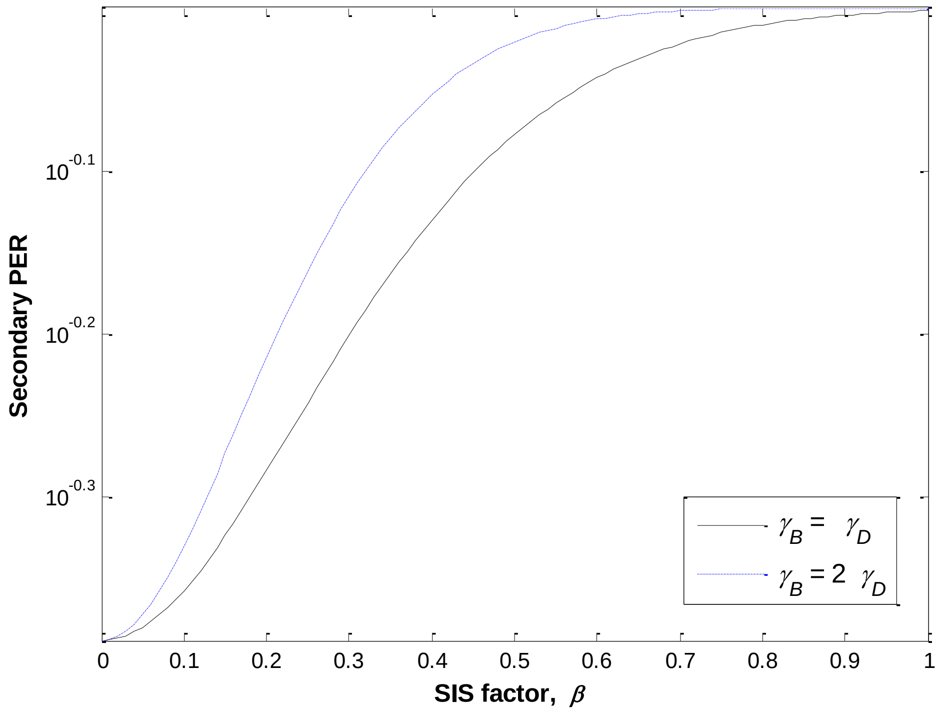


Fig. 3.16. Secondary PER against SIS factor

And finally Fig. 3.17 and Fig 3.18 show the normalized secondary throughput versus secondary SNR for equal and unequal primary and secondary power settings. As we observe from (3.25), the secondary network throughput is affected not only by the packet error rate of the primary network, but also by the amount of residual interference pertaining to full duplex operation. When there are more packet errors in the primary network, more opportunities for secondary transmission emerge. On the other hand, secondary throughput is inversely related to its own packet error rate. Therefore, a maximum is expected when the SNR of both networks varies. As it is seen, for the scenario of simulation, this maximum occurs around 10 dB when primary and secondary powers are equal, and around 12 dB when secondary power is half of the primary power. It is obvious that imperfect SIS has a negative effect on secondary throughput which is verified in these figures.

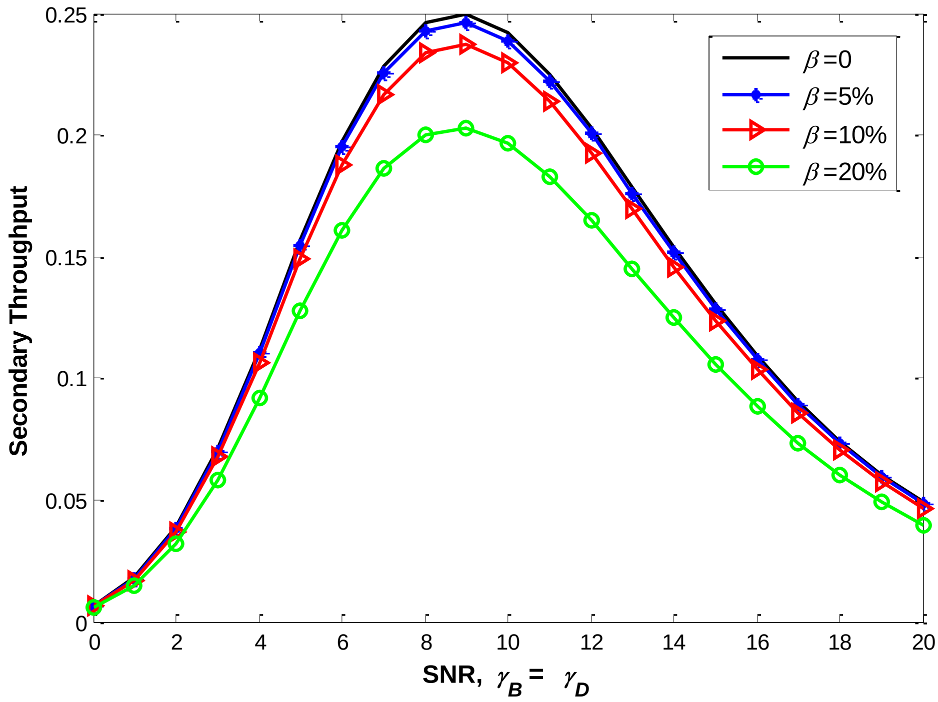


Fig. 3.17. Secondary throughput for equal primary and secondary powers.

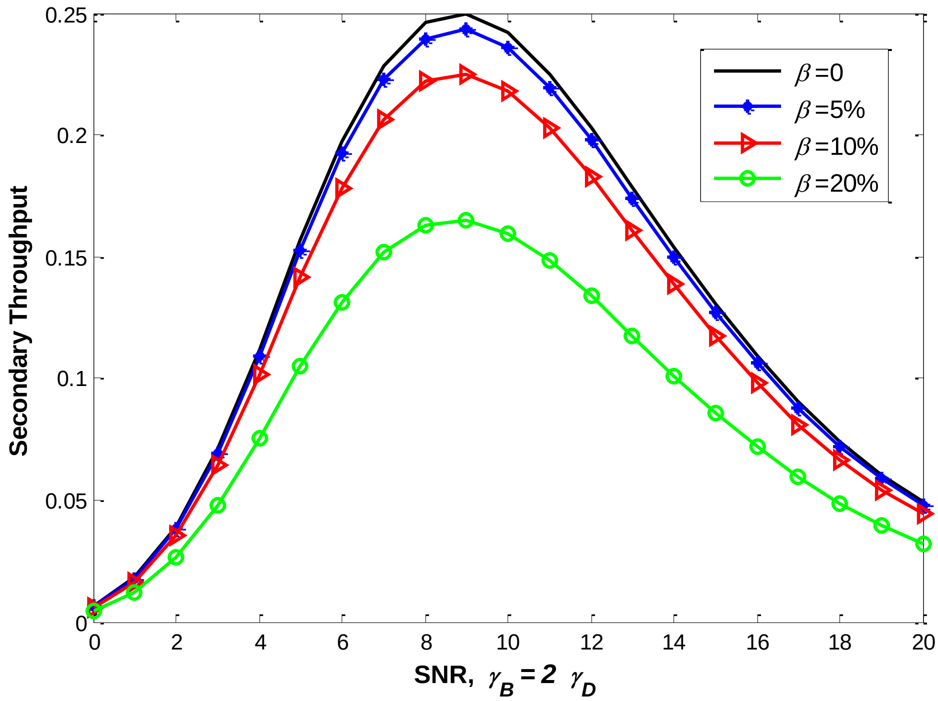


Fig. 3.18. Secondary throughput for unequal primary and secondary powers.

3.6 Summary

In this chapter we proposed a full duplex cooperative retransmission protocol for an overlay cognitive radio network with ARQ retransmissions. In the proposed system the primary network is oblivious to the presence of the secondary network, but will benefit from its cooperation whenever the SU takes part.

During ARQ retransmission rounds, the secondary network exploits the redundancy introduced in the system to transmit data, and simultaneously cooperates with the primary network in repeating the failed packet in a full duplex manner in order to compensate for inflicted interferences. This requires self-interference cancellation as well as known interference cancellation capabilities of the destination node in the secondary network, which are achievable with recent advances in KIC and SIS technologies.

The primary network may employ simple ARQ or Hybrid ARQ schemes for detection of failed packets, and performance of both methods in the proposed scheme has been investigated. Packet error rate as a practical metric in packet-based design paradigms, and expected throughput of primary and secondary networks for ARQ and HARQ scenarios were mathematically analyzed and numerically verified through simulations.

We demonstrated that the proposed protocol improves throughput of both secondary and the primary networks compared to other non-cooperative and half duplex protocols and has less complexity for implementation, in contrast to the methods based on network coding [107], [108].

Chapter 4

Asynchronous Full-Duplex Cognitive Radio

4.1 Introduction

Although the first applications of FD technology dates back to half a century ago, it has not been widely used until recently, due to strong self-interference from own transmitter. Fortunately, new advances in self-interference suppression methods have opened new doors for practical application of FD communication in general mobile systems.

As the main motivation behind introduction and development of CR technology was the optimum exploitation of the wireless spectrum, any idea which can improve the spectrum utilization in a CRN would be interesting. Application of FD technology in a CRN is one of those ideas which has attracted much attention recently and seems to be the technology of future wireless communications.

As mentioned in Chapter 2, full duplex technology has other advantages than doubling the spectral efficiency, which may be beneficial in a cognitive radio system.

In this chapter we have proposed a novel scheme in an interweave CR network, where asynchronous FD transmission in the secondary network results in enhanced spectral efficiency as well as improvements in performance of the whole system through reduction in probability of collision and average collision duration. Cooperative sensing is another feature of the proposed scheme which is described in detail in this chapter.

4.2. Literature Review

As a new field of research, there are not many studies on full duplex cognitive radio networks. This is due to the fact that full duplex communication was not deemed possible until recently. Most studies on full duplex cognitive radio have concentrated on simultaneous transmitting and sensing [55]-[59] which does not improve spectrum utilization directly, and mostly enhances the network performance in other terms. Some studies [55]-[58] have proposed full duplex spectrum sensing to improve the primary network protection against collisions with the secondary transmission. In such works the transmitting SU keeps sensing the presence of PUs constantly during the transmission slot, and upon detection of a PU signal, will stop transmission to avoid collision [55]. This technique requires a near perfect self-interference cancellation capability for secondary nodes if conventional sensing methods based on energy detection are involved. Although such techniques will noticeably improve PU protection, they will not enhance throughput of the CR network. However, the main advantage of using FD technology is doubling the systems throughput. The authors in [88] presented a threshold-based sensing-transmission scheme in order to maximize SU throughput with minimum collision with PUs in a half-duplex network, which may be applied in full duplex mode either. Their scheme consisted of multiple consecutive sensing or transmission periods, determined according to the SU's belief about the state of the PU(s). In their scheme, the SU has only two options: spectrum sensing or data transmission. The objective was to maximize the SU's utility, which rewards the SU for successful transmission and penalizes it for collisions.

Another adaptive scheme was proposed in [89], where a secondary transmitter adapts its sensing and transmission durations according to its belief regarding the PU state of activity. In this case, the SU can either stay idle, sense the spectrum, or transmit its data. Although these schemes show enhancement over traditional LBT schemes in terms of throughput and collision probability in half duplex mode, in full duplex mode they would not be optimal. Inspired by that idea, the authors in [89] proposed a similar approach for a FD scenario in a CR network with four modes of operation; sensing-only, transmit and sensing, transmit and receive, and channel selection, and developed an optimal mode-selection strategy to maximize the SU throughput for a given PU collision probability. The criterion for choosing the optimal action was to maximize the SU's utility subject to a constraint on the PU collision probability. They considered the waveform-based spectrum sensing technique for the transmit and sensing mode, in order to overcome detection inaccuracies at imperfect SIS inherent to energy detection methods.

In this chapter we have considered exploitation of in-band full duplex transmission and reception in an interweave cognitive radio network where secondary users are capable of imperfect self-interference suppression. We have presented a novel scheme for full duplex CR networks based on two modes of operation; Cooperative Sensing (CS) mode, and Full Duplex Transmit and Receive (FDTR) mode.

The benefits of cooperative sensing in half duplex CR networks have been well studied in many works [90]-[92] but have not been fully investigated in bidirectional full duplex schemes so far. When a primary user is active, secondary users keep sensing the channel through consecutive sensing periods in a cooperative manner. This will increase detection probability and there would be no collision with the primary network during its busy period. In FDTR mode, we have no regular sensing periods and detection of PU's activity is implemented through feedback information (ACK/NACK) from SUs' receivers. In this mode, collision with the PU is inevitable, but we have considered asynchronous transmission of SUs as a new approach to decrease collision duration compared to other methods. To the best of our knowledge, the benefits of asynchronous FD communication in a CR network has not been studied, and only its effects on enhancing self-suppression capability have been investigated [93]. We have shown that such alteration in transmission timing in the secondary network, will decrease collision duration to half compared to that in the traditional sensing methods and have found the optimum timing difference for minimum collision duration.

The main contributions of this chapter are as follows. We have presented a novel scheme for full duplex operation in cognitive radio networks which consists of cooperative sensing and asynchronous full duplex transmission by secondary users. The proposed scheme operates in two modes; the cooperative sensing (CS), and full duplex transmit and receive (FDTR) mode. We have analyzed the impact of these methods on network metrics such as probability of false alarm and probability of detection, probability of collision and average collision duration, and the CR network average throughput. We have derived closed forms for probability density function of the collision duration in a CR network where primary users' activity is modeled by an ON/OFF Markov process, and channels between primary and secondary nodes are Rayleigh flat fading. In addition, we have derived analytical closed forms for average throughput of the secondary network under such conditions and have shown that our scheme outperforms the synchronous and traditional methods in terms of cognitive network metrics.

The rest of this chapter is organized as follows. Section 4.3 describes the system model. In Section 4.4 we derive exact closed form expressions for probability of detection and false alarm in the CS and FDTR modes, probability of collision and probability density function of collision duration, average collision duration and cumulative collision duration, and the average CR network throughput for the FDTR mode (synchronous and asynchronous) and the traditional LBT scheme for comparison. Numerical results are provided in Section 4.5, and the summary and concluding remarks are presented in Section 4.6. We have summarized the major notations used in this chapter in Table 4.1.

TABLE 4.1. Significant Notations.

P1, P2, S1, S2	Primary and secondary nodes, resp.
β	SIS factor
μ^{-1} , λ^{-1}	Primary network average ON and OFF times, resp.
T_1, T_2, T_s	S1 frame, S2 frame, and sensing slot durations, resp.
P_{f_1}, P_{d_1}	Probability of false alarm and detection in CS mode, resp.
P_{f_2}, P_{d_2}	Probability of false alarm and detection in FDTR mode, resp.
α, g, γ_t	PER fitting parameters
γ_{S1}, γ_{S2}	SNR of received signal at S1 and S2, resp.
τ, θ	Duration of collision and non-collision, resp.
P_τ, P_{out}	Probability of collision and outage, resp.
R_0, R_1	Secondary network nominal and degraded throughput, resp.
$R_{Sync}, R_{Async}, R_{LBT}$	Sync., async., and LBT modes throughput, resp.
$P_{Async}^{col}, P_{LBT}^{col}$	Probability of collision in async. and LBT modes, resp.

4.3 System Model

We consider an interweave cognitive radio network, consisting of multiple primary users and one pair of secondary users, which opportunistically access the licensed PUs' channel as shown in Fig. 4.1. We focus on a bidirectional channel, and assume all channels between primary to secondary, and secondary to secondary nodes are quasi-static Rayleigh block fading. Thus, the channel is constant within a coherence interval and changes independently from one coherence interval to another. PUs may become active or inactive at any time, modeled as an alternating ON/OFF continuous-time Markov process which has been used in many works [101], [102]. The ON and OFF times are exponentially distributed with mean durations of μ^{-1} and λ^{-1} respectively. The secondary users are capable of operating in the in-band full duplex mode i.e. to transmit and receive at the same time over the same frequency band. This requires each secondary device be capable of perfect or partial self-interference suppression. Here we do not consider SIS methods and similar to the previous chapter, just quantify this capability by SIS factor β_i of the ith SU. SIS factor is the ratio between the residual self-interference signal and the original one. In perfect SIS (i.e. no residual self-interference) β_i=0; and otherwise, only the fraction $1-\beta_i$ of the original self-interference signal is cancelled. In this thesis we assume both SUs have the same SIS factor β.

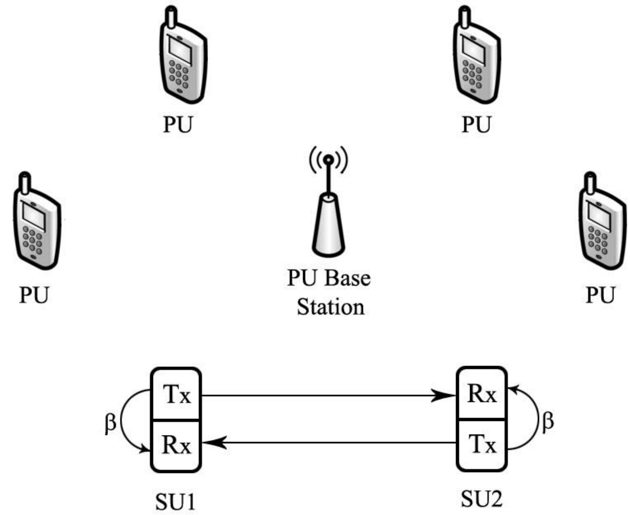


Fig. 4.1. System model of our network comprising of a primary base station, some primary receivers, and a pair of secondary users capable of FD operation.

In order to avoid collision with primary users, the SUs sense the channel cooperatively and if it is found idle, will start bidirectional transmission. When a PU returns to the channel, the SUs will be informed through collision and will stop transmission instantly. Energy detection is the core of channel sensing in this method. Since the sensing is done in the half duplex mode, the complications of imperfect SIS will not affect the detection procedure. The SUs operate in two modes; Cooperative Sensing (CS) Mode, and Full Duplex Transmit and Receive (FDTR) mode.

In the CS mode, both SUs just listen to the channel for T_s seconds (sensing slot), which is a fraction of secondary transmission frame duration T, and do not transmit. This is similar to the traditional cooperative sensing in half duplex networks [90]. If both SUs detect a spectrum hole, then both will switch to the FDTR mode. Cooperative sensing is carried out through a handshaking MAC protocol as follows.

4.3.1 Cooperative Sensing MAC

Each secondary senses the channel through energy detection methods to be informed of the presence or absence of primary users. When one SU finds the channel not in use, it transmits RTS (ready to send) signals (which are very short and assumed to be error free) in the next sensing slot and at the same time listens to receive a RTS signal back from the peer SU. Any SU that detects a signal within a sensing slot, first try to decode an RTS out of it. If successful, it will emit an RTS back and will start transmission and reception in the FDTR mode accordingly. It is obvious that this implies that the other peer has the same conditions and will go to the FDTR mode either. But if one SU sends RTS signals but does not receive an RTS back, or it cannot decode an RTS from the signal detected within sensing slots, it means that the other end has not found the channel idle, or a PU is present, so the sensing continues in the following intervals until both arrive at the same decision. This process is shown in Fig. 4.2.

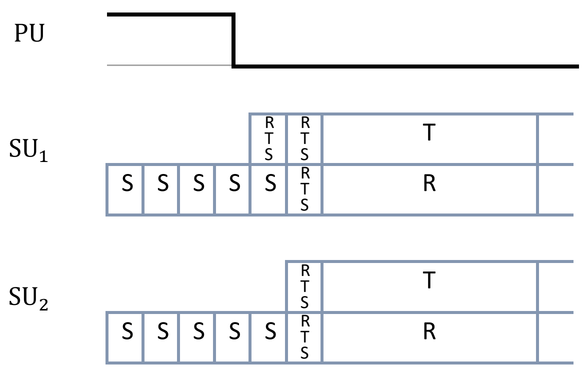


Fig. 4.2. Cooperative sensing MAC in CS mode.

In FDTR mode, SU1 transmits its packets within a frame with size T_1 and SU2 transmits within a frame with size T_2, and as long as they have not detected or been informed of the return of a primary signal, will remain in this mode of operation. Detection of a PU signal in the FDTR mode is achieved through collision event. It will be seen later that if T_1 and T_2 are set unequal (which implies asynchronosity of transmissions), duration of the collision with the PU would be less than in the case of synchronous transmission with equal frame lengths (i.e., $T_1 = T_2$). In our proposed

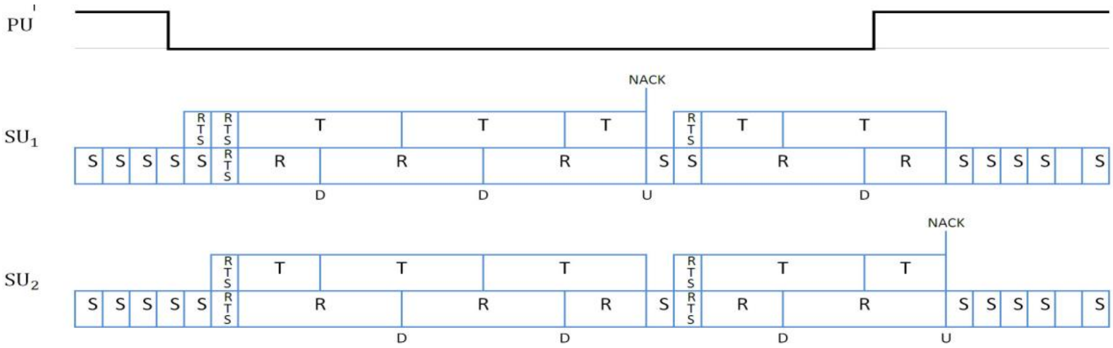


Fig. 4.3. CS and FDTR modes and switching from CS mode to FDTR mode triggerred by NACK and Undecode (U) events caused by frame error (false alarm) or collision with primary.

scheme we consider the case when $T_1 = T_2 = T$ and apply a new idea to decrease the collision duration.

When the primary signal reappears, it will collide with both SUs' transmissions and this will end in SU packet collision and frame error. Any SU that cannot decode the received packet or frame without error (i.e. Undecode event), or receives a NACK from the other SU (i.e. NACK event), would stop transmission immediately and both units will go into the CS mode again. The whole process in the CS and FDTR modes is shown in Fig. 4.3.

Although operating in the FDTR mode enhances the SUs' throughput, the SUs will not be able to monitor the PUs' state within a transmission frame. Hence, the probability of colliding with the PU will be higher than in traditional methods of sensing, such as Listen Before Talk (LBT). In addition, the PUs' signal may affect SUs differently due to different fading or interference conditions of PU-SU channels. It means that a SU transmission may not be affected by the PU signal, due to deep fading of the channel between the primary transmitter and the secondary receiver, and no packet or frame error would be declared although PU is present. This may increase collision probability as well. But such conditions may happen during sensing slots in traditional methods too, and could result in longer collision durations.

This method has the drawback that there may be packet or frame errors not caused by collision with a primary signal. Fading, noise or any other interferences may cause such errors which will in turn cause false alarm, interrupt SUs' transmissions, and decrease SU throughput. However, our method is a conservative way to decrease the duration of any probable collision with the primary users, which guarantees high primary protection. On the other hand, duration of sensing slots is small compared to the secondary frames, and cooperative signal detection decreases the probability of consecutive false alarms. Therefore, the throughput degradation due to general packet or frame errors is small, as will be seen later in the analysis and numerical results.

4.3.2 Asynchronous Full Duplex

The sensing method applied in the FDTR mode which is based on collision event usually results in long collision durations and may not be tolerated by the primary network. In order to alleviate this, we propose a simple change in full duplex synchronization. If both SUs are synchronous and transmit at the same time, upon collision with a PU signal, both NACK signals would be generated at the same time, at the end of the frame. But if SUs transmit asynchronously (or have different durations), then these detection events would not be coincident and the result of it would be shorter average collision durations as depicted in Fig. 4.4. For the case of equal transmission frames $T_1 = T_2 = T$, the optimum timing difference would be $T/2$ for shortest average collision duration.

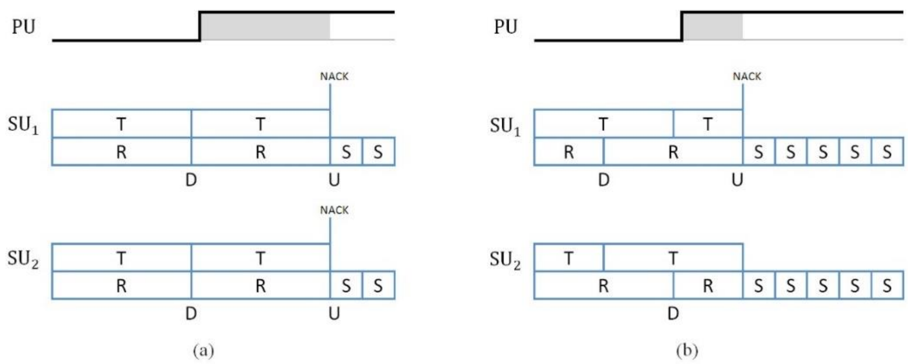


Fig. 4.4. a) Synchronous FD transmission, and b) asynchronous FD transmission

Proposition: For the case of equal transmission frames $T_1 = T_2 = T$, the minimum upper limit of a collision duration happens when timing difference is $L = T/2$, and the minimum average collision duration occurs when timing difference is $L = \frac{1+\lambda T - W(e^{1+\lambda T})}{\lambda}$ in which $W(\cdot)$ is the Lambert W function.

Proof: As it is shown in Fig. 4.5. for a time difference of L seconds, the duration of collision in a frame would be:

$$\tau = \begin{cases} L-\theta & if \quad 0<\theta<L \\ T-\theta & if \quad L<\theta<T \end{cases}$$

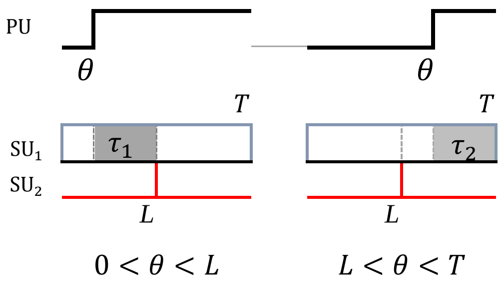


Fig. 4.5. Duration of collision

in which random variable θ with pdf $p(\theta) = \frac{\lambda e^{-\lambda\theta}}{1-e^{-\lambda T}}$ indicates the end time of primary's OFF period. For $0 < \theta < L$ the maximum duration will occur at $\theta = 0$ and $\tau_1 = L$. For $L < \theta < T$ it occurs at $\theta = L$ and $\tau_2 = T - L$. The minimum would happen when $\tau_1 = \tau_2$which will result $L_{min} = T/2$.

The average collision duration in a single frame would be:

$$\begin{aligned}
\bar{\tau} &= P(0<\theta<L)(L-\theta) + P(L<\theta<T)(T-\theta) \\
&= \int_0^L p(\theta)(L-\theta)d\theta + \int_L^T p(\theta)(T-\theta)d\theta \\
&= L\int_0^L p(\theta)\,d\theta + T\int_L^T p(\theta)\,d\theta - \int_0^T p(\theta)\,d\theta \\
&= L\left(\frac{1-e^{-\lambda L}}{1-e^{-\lambda T}}\right) + T\left(\frac{e^{-\lambda L}-e^{-\lambda T}}{1-e^{-\lambda T}}\right) + \left(\frac{(\lambda T+1)e^{-\lambda T}-1}{\lambda(1-e^{-\lambda T})}\right)
\end{aligned}$$

In order to find the optimum value for L which minimizes above equation, we have to form its derivative with regard to L and set it to zero:

$$1+(\lambda L-1)e^{-\lambda L}-\lambda T e^{-\lambda L}=0$$

$$(\lambda L-1-\lambda T)e^{-\lambda L}=-1$$

$$(-\lambda L+1+\lambda T)e^{(-\lambda L+1+\lambda T)}=e^{1+\lambda T}$$

$$-\lambda L+1+\lambda T=W\left(e^{1+\lambda T}\right)$$

$$L_{opt}=\frac{1+\lambda T-W(e^{1+\lambda T})}{\lambda}$$

where $W(.)$ is the Lambert W function. □

4.4 Performance Analysis

4.4.1 Sensing metrics

If prior knowledge of the PU signal is not available, the energy detection method is optimal for detecting zero-mean constellation signals in half duplex paradigms [94]. In our proposed system, sensing in CS mode is done in a half duplex manner. In order to increase detection probability, cooperative sensing is performed through consecutive sensing slots of duration T_s, sampling frequency f_s and with $N=f_s\cdot T_s$ samples in each slot. For circularly symmetric complex Gaussian primary and noise signals, the probabilities of false alarm, $P_f(\epsilon,T_s)$ and detection, $P_d(\epsilon,T_s)$, for each node are presented in (2.1) and (2.2) respectively, and probability of misdetection, $P_m(\epsilon,T_s)$ would be:

$$P_m(\epsilon,T_s)=1-P_d(\epsilon,T_s), \tag{4.1}$$

For a lower probability of collision, we will fuse the results of two sensing units and the hypothesis is that PU is in off state when both SUs have sensed the channel not in use. This will decrease misdetection probability to a very low value but will increase the probability of false alarm [94]:

$$P_{f1} = 2P_f(\epsilon, T_s) - P_f(\epsilon, T_s)^2 \;,\; P_{d1} = 2P_d(\epsilon, T_s) - P_d(\epsilon, T_s)^2. \quad (4.2)$$

P_{f1} and P_{d1} are probabilities of false alarm and detection in the CS mode. In the FDTR mode detection method is different and is based on NACK and Undecode events when a collision with primary happens. Since we assume that transmission errors in NACKs are negligible, the probability of not receiving a NACK at a secondary node (misdetection) despite the presence of primary signal, is the probability that both channels from primary transmitter node to SU1 and SU2 be in deep fade and the reverse channels (i.e. from SU1 and SU2 to primary receiver) are not in deep fade, which is very small and may be ignored. On the other hand, false alarm happens when an error in detection of a frame triggers a NACK due to reasons other than collision with the primary. Hence within the FDTR mode, false alarm probability (P_{f2}) is the same as the average Frame Error Rate ($\overline{FER}$) [1]. We assume that both SUs have the same transmission parameters and average frame error rate. Average frame error rate for transmissions with N_f packets in a frame would be:

$$\overline{FER} = 1-(1 - \overline{PER})^{N_f} \;. \quad (4.3)$$

$\overline{PER}$ is the average packet error rate of the secondary network. Here, similar to previous chapter, we use the approximate value for instantaneous packet error rate given in [87] which is widely used in many works:

$$PER(\gamma) = \begin{cases} 1, & 0 < \gamma < \gamma_t \\ \alpha \exp(-g\gamma)\,, & \gamma \geq \gamma_t \end{cases} \quad (4.4)$$

where γ is the instantaneous signal-to-noise ratio and $(\alpha, g,\, \gamma_t)$ are mode dependent parameters found by least-squares fitting to the exact packet error rate. In a Rayleigh fading channel, γ follows an Exponential distribution with the following probability density function:

[1] If NACK and Undecode events are based on packet errors instead of frame errors, then $P_{f2} = \overline{PER}$.

$$f_\gamma(\gamma) = \frac{1}{\bar{\gamma}} \exp\left(\frac{-\gamma}{\bar{\gamma}}\right), \tag{4.5}$$

where $\bar{\gamma}$ is the average received signal-to-noise ratio. For SU2 with imperfect SIS, the overall instantaneous SNR at the receiver is $\gamma_{S_1}/(1+\beta\gamma_{S_2})$ in which $\gamma_{S_1}$is the SNR of SU1 at SU2 and $\beta\gamma_{S_2}$is the residual self-interference of SU2 after partial SIS. Therefore, the average value of packet error rate would be:

$$\begin{aligned} \overline{PER} = \mathbb{E}[PER(\gamma)] &= \int_0^\infty PER\left(\frac{\gamma_{S_1}}{1+\beta\gamma_{S_2}}\right) f_{S_1}(\gamma_{S_1})\, d\gamma_{S_1} \\ &= 1 - \frac{g\bar{\gamma}_{S_1}}{1+\beta\bar{\gamma}_{S_2}+g\bar{\gamma}_{S_1}} \exp\left(\frac{-(1+\beta\bar{\gamma}_{S_2})\gamma_t}{\bar{\gamma}_{S_1}}\right) \end{aligned} \tag{4.6}$$

As it is seen, the residual self-interference factor (β) and frame length are the factors which affect the false alarm probability in the FDTR mode which will be investigated numerically in Section 4.4.

4.4.2 Probability of Collision and Collision Duration Distribution

As depicted in Fig. 4.6, collision with the primary signal may occur at the beginning (pre-collision) or the end (post-collision) of a secondary frame. Pre-collision happens when the network is in the CS mode and misdetection occurs, and post-collision occurs within the FDTR mode, when a primary signal reappears to the channel. Therefore, probability of collision, P_{Async}^{col} in the asynchronous scheme would be

$$P_{Async}^{col} = P(\mathcal{H}_0)\left(1 - P_{f_2}\frac{T_S}{T}\right)P_\tau + P(\mathcal{H}_1)(1-P_{d1}), \tag{4.7}$$

where $P_\tau = 1 - e^{-\lambda T}$ is the probability that primary signal reappears within a SU frame of duration T. In the traditional LBT scheme, probability of collision can be written as:

$$P_{LBT}^{col} = P(\mathcal{H}_0)\big(1-P_{f1}\big)P_\tau + P(\mathcal{H}_1)(1-P_{d1}). \tag{4.8}$$

$\mathcal{H}_0$ is the event that the PU is not active at the start of the frame, and its complement is $\mathcal{H}_1$. In our system where PU's idle and busy times are Exponential with parameters λ and μ respectively, the respective *a priori* probabilities of PU activity at the start of each secondary frame are $P(\mathcal{H}_0) = \frac{\mu}{\lambda+\mu}$ and $P(\mathcal{H}_1) = \frac{\lambda}{\lambda+\mu}$.

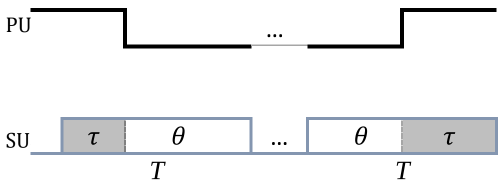


Fig. 4.6. Pre and post collision events

When a collision happens in the FDTR mode, it would last until the first NACK signal is declared from any of the SUs. Having assumed that the NACK signals are free from errors, in asynchronous transmission mode when there is a difference of $T/2$ seconds between the SUs' transmissions timing, the probability density function (pdf) of duration of collision (τ) would be as follows:

$$p_\tau(\tau) = \begin{cases} \dfrac{\lambda e^{-\lambda(\frac{T}{2}-\tau)}}{1-e^{-\lambda\frac{T}{2}}} & , \quad 0 < \tau < \dfrac{T}{2} \\ 0 & , \quad \tau > \dfrac{T}{2} \end{cases} . \qquad (4.9)$$

Collision duration may be longer than $T/2$ if there is no NACK signal at the end of a transmission frame. This may happen when the channel between the SU transmitter to at least one of the PU receivers has been in deep fade for at least the entire collision duration. If channels from a PU receiver to both SU receivers have been in deep fade since PU has reappeared and collided with SU signals, then we can conclude that SUs' transmissions have no harm on the PU, and in fact no collision has occurred although the PU is present. In order to derive probability density function of collisions with durations more than $T/2$, we need to calculate the minimum duration outage probability. In a Rayleigh flat fading channel, the probability of minimum duration outage $P_{out}(\tau)$ is the probability of *SIR* going below a required threshold *R* and staying there for at least τ seconds. There have been many studies on the minimum duration outage probability (or briefly outage probability) mostly based on Gamma function or Weibull distribution [95]-[98]. Authors of [97] have proposed an analytical approximation for minimum duration outage probability in a Rayleigh channel, based on asymptotic result for the fade duration distribution function of deep fades given by Rice [98]. Here we use the approximation for the probability of minimum duration outage presented in [97]:

$$P_{out}(\tau) = f_{out}(\tau).\bar{\tau}_{out}(\tau), \tag{4.10}$$

$$\bar{\tau}_{out}(\tau) = \tau + \frac{\tau_m}{2I_1\left(\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right) exp\left(-\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right)} \cdot \sum_{k=0}^{\infty} \frac{1}{k!\,(k+1)!\,2^{2k+1}}$$
$$\cdot\left\{(2k)! - \left[\left(\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right)^{2k} + (2k)\left(\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right)^{2k-1} + \cdots + (2k)!\right] exp\left(-\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right)\right\}, \tag{4.11}$$

$$f_{out}(\tau) = 2\sqrt{2\pi f_m \rho exp(-\rho^2)\left(\frac{\bar{\tau}}{\tau}\right) I_1\left(\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right)} \cdot exp\left(-\frac{2}{\pi}\left(\frac{\bar{\tau}}{\tau}\right)^2\right), \tag{4.12}$$

where f_m is the maximum Doppler frequency of the channel, $\rho = R/R_{rms}$ is the normalized threshold, $\bar{\tau}$ is the average fade duration in a Rayleigh channel, and I_1 is the modified Bessel function of the first kind.

Hence, collisions with durations more than $T/2$ will happen when we have an outage at the same time on one of the channels. The respective probability density function of collision duration can be derived from:

$$p_\tau(\tau) = \begin{cases} \dfrac{\lambda e^{-\lambda\left(\frac{T}{2}-\tau\right)}}{1-e^{-\lambda\frac{T}{2}}}(1-P_{out}(\tau)) & , \quad 0 < \tau < \dfrac{T}{2} \\ \dfrac{\lambda e^{-\lambda(T-\tau)}}{1-e^{-\lambda\frac{T}{2}}} P_{out}(\tau) & , \quad \dfrac{T}{2} < \tau < T \\ 0 & \quad \tau > T \end{cases} \tag{4.13}$$

The probability that collision takes longer than T seconds would be so small, hence we will ignore it.

4.4.3 Throughput

When a PU is active, our proposed system is operating in the CS mode, sensing the spectrum, and none of the SUs transmits. However, in case of a misdetection, they may start transmitting while the PU is on. Since the probability of misdetection in the CS mode is very low, we ignore such rare possibility. On the other hand, when the PU is idle and the system is in the FDTR mode with imperfect SIS capability, SUs transmit in a full-duplex manner with total throughput $R_0 = 2\log(1 + \frac{SNR_S}{1+\beta SNR_S})$ during collision free intervals, and with degraded throughput $R_1 = 2\log(1 + \frac{SNR_S}{1+SNR_p+\beta SNR_S})$ during collision interval (SNR_S is the signal-to-noise ratio of the SUs, and SNR_p is that of PU at SUs' receiver). If any false alarm is announced in between, throughput will be zero during the subsequent sensing interval T_s. Referring to Fig. 4.6, we denote the transmission time without collision in a frame by θ, and the collision duration by τ. Then the throughput of a frame in the transmission phase would be:

$$R_T = P_\tau \left(R_0 \frac{\theta}{T} + R_1 \frac{\tau}{T} \right) + R_0(1 - P_\tau), \tag{4.14}$$

and the average value of the throughput over a frame would be:

$$\bar{R}_T = P_\tau \left(R_0 \frac{\bar{\theta}}{T} + R_1 \frac{\bar{\tau}}{T} \right) + R_0(1 - P_\tau), \tag{4.15}$$

where $\bar{\tau}$ and $\bar{\theta}$ are, respectively, the average values of non-collision and collision durations in asynchronous FDTR mode and are derived as follows:

$$\bar{\tau} = \mathbb{E}\left[\tau | \tau < \frac{T}{2}\right] = \int_0^{T/2} \tau p_\tau \left(\tau | \tau < \frac{T}{2}\right) d\tau = \int_0^{T/2} \tau \frac{\lambda e^{-\lambda\left(\frac{T}{2}-\tau\right)}}{1 - e^{-\lambda\frac{T}{2}}} d\tau = \frac{\lambda T + e^{-\lambda T} - 1}{2\lambda(1 - e^{-\lambda T})}, \tag{4.16}$$

$$\bar{\theta} = \mathbb{E}[\theta | \theta < T/2] = T - \bar{\tau} = \frac{1 + \lambda T - (2\lambda T + 1)e^{-\lambda T}}{2\lambda(1 - e^{-\lambda T})}. \tag{4.17}$$

The first two terms in parenthesis in (4.15) correspond to the case when the secondary frame encounters collision, and the second term corresponds to the frame with no collision. In order to calculate the average throughput, we need to consider the effect of false alarms on it. As we have two modes of operation and two types of false alarms corresponding to each mode, we need to differentiate between them. False alarm effects on the first secondary frame following the CS mode differently from other frames within the FDTR mode. When the primary signal disappears, the CR network would detect an opportunity with probability of $1-P_{f1}$ and will start transmission for at least half of a frame duration, or will miss the opportunity for T_s seconds (sensing slot) with probability of P_{f1}. Within the FDTR mode, if a false alarm occurs (with probability of P_{f2}), the system will go into the CS mode for at least one sensing slot (i.e. T_s seconds), otherwise will continue transmission for $T/2$ seconds by the end of which another ACK/NACK signal is issued. Given the above breakdown, we derive the average throughput for a bidirectional asynchronous full duplex CR, based upon our proposed scheme as follows:

$$R_{avg} = 2P(\mathcal{H}_0)\left[\frac{T}{\lambda}\left(1-(P_{f1}+P_{f1}^2+\cdots)\frac{2T_s}{T}\right)\bar{R}_T + (1 - \frac{T}{\lambda})\left(1-(P_{f2}+P_{f2}P_{f1}+P_{f2}P_{f1}^2+\cdots)\frac{2T_s}{T}\right)\bar{R}_T\right], \tag{4.18}$$

in which $\frac{T}{\lambda}$ corresponds to the probability that the considered frame is the first frame following the CS mode, and $(1-\frac{T}{\lambda})$ accounts for the frames in the FDTR mode. The terms $(P_{f1}+P_{f1}^2+\cdots)$ and $(P_{f2}+P_{f2}P_{f1}+P_{f2}P_{f1}^2+\cdots)$ in (4.21) account for consecutive false alarms in the CS and FDTR modes, respectively, and may be truncated to only the first terms with a good approximation. Hence we can rewrite the above equation as follows:

$$R_{avg} \cong 2P(\mathcal{H}_0)\left(1+\frac{2T_s}{\lambda}\left(P_{f2}-P_{f1}\right)-\frac{2T_s}{T}P_{f2}\right)\{P_\tau(R_0\bar{\theta}+R_1\bar{\tau})+R_0(1-P_\tau)\}. \tag{4.19}$$

Having inserted the respective values for $\bar{\tau}$ and $\bar{\theta}$ from (4.16) and (4.17) in (4.19) and after some manipulation, the average throughput of a CR network with the proposed FD scheme, in asynchronous and synchronous transmission modes, respectively are derived as:

$$R_{Async.} = 2\,P(\mathcal{H}_0)\left(1+\frac{2T_s}{\lambda}\left(P_{f2}-P_{f1}\right)-\frac{2T_s}{T}P_{f2}\right)\left[(R_0-R_1)\frac{1-\lambda T-e^{-\lambda T}}{2\lambda T}+R_0\right], \tag{4.20}$$

$$R_{Sync.} = 2P(\mathcal{H}_0)\left(1+\frac{2T_s}{\lambda}\left(P_{f2}-P_{f1}\right)-\frac{2T_s}{T}P_{f2}\right)\left[(R_0-R_1)\frac{1-\lambda T-e^{-\lambda T}}{\lambda T}+R_0\right]. \tag{4.21}$$

For comparison purposes we have derived the throughput of a full duplex cognitive system using the LBT technique with similar duration of sensing slots T_s and transmission frames T, as follows:

$$R_{LBT} = 2P(\mathcal{H}_0)\left(1 - P_{f1}\right)\left[R_1\left(1 - \frac{T_s}{T}\right) + (R_0 - R_1)\frac{1 - e^{-\lambda(T - T_s)}}{\lambda T}\right]. \tag{4.22}$$

4.5 Simulation Results

In this section, we present some simulation results to demonstrate the performance of the proposed schemes. We have considered a CR network with a pair of SUs and a pair of PUs. The parameters and characteristics which are utilized in the simulations are as follows: Activity parameters of the PU network; $\lambda^{-1} = 150$ msec (OFF state) and $\mu^{-1} = 300$ msec (ON state), sampling rate f_s= 1MHz, BPSK modulation with convolutional coding, packet size= 1080 bits, CR frame size= 2 packets, secondarys' SNR= 10 dB, primary SNR at SUs =3 dB, and size of sensing slot $T_s = 1$ ms. We have used the parameters presented in Tab. 3.2 and Tab. 3.3 for calculating packer error rate of SUs' communication.

Firstly, we study the performance of the sensing metric of the proposed scheme in the CS mode and the FDTR mode. Fig. 4.7 shows the impact of duration of sensing slot T_s on the probability of false alarm and detection for cooperative sensing (CS) mode and traditional non-cooperative sensing mode, as in LBT systems. As it is seen detection probability in cooperative mode is more than that of single non-cooperative sensing and quite close to one, and its changes vs. sensing duration is negligible. On the other hand, false alarm probability in cooperative mode is more than single mode which may appear as a drawback. But as will be seen later, this slight difference has very little impact on CR network performance metrics thanks to the scheme applied.

Fig. 4.8 shows the false alarm probability in the FDTR mode vs. SIS factor for different frame size. As we see both frame size and SIS factor directly affect false alarm probability in the FDTR mode according to (4.3) and (4.6). When there is some residual

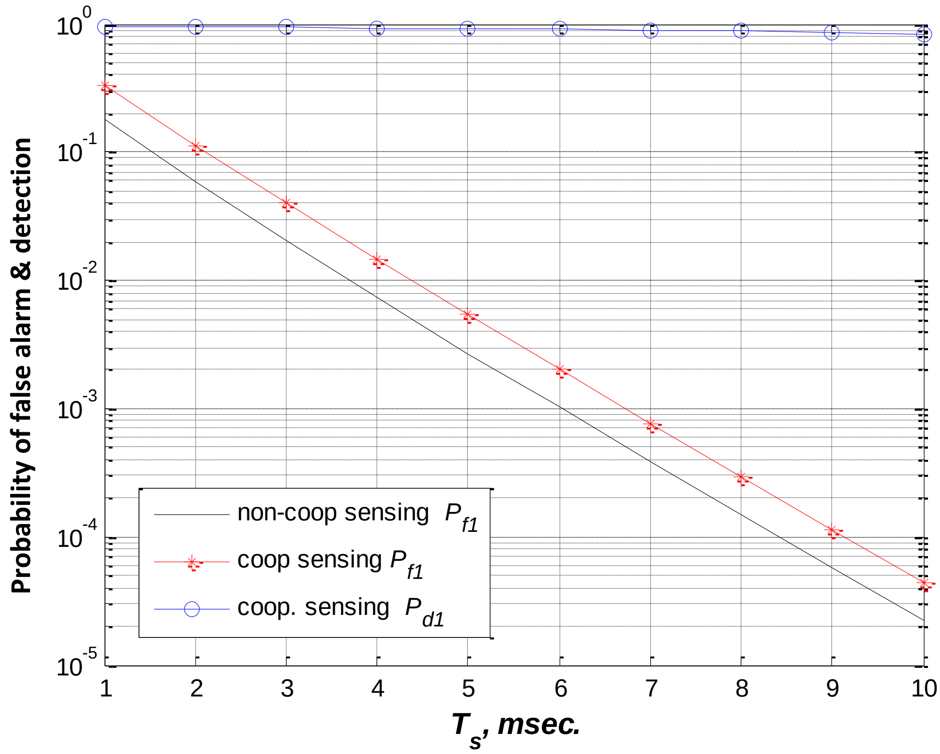


Fig. 4.7. Probability of false alarm (P_{f1}) and detection (P_{d1}) vs. sensing time duration T_s for cooperative and non-cooperative modes.

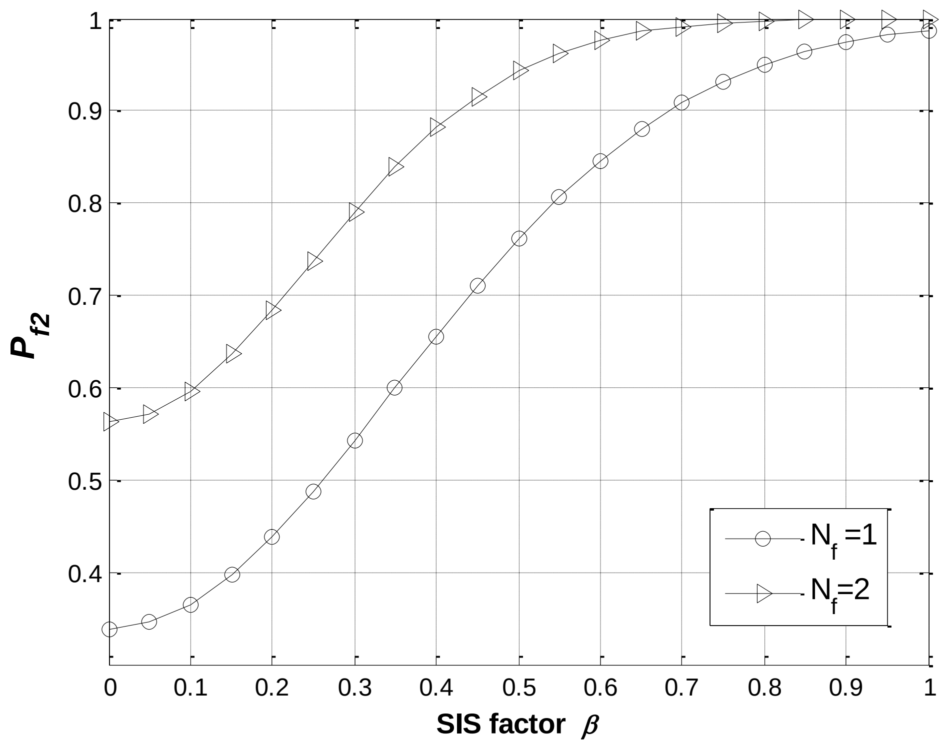


Fig. 4.8. Probability of false alarm (P_{f2}) vs. SIS factor (β) in FDTR mode for different frame sizes

the false alarm probability will increase. Number of packets in a frame (N_f) has a similar effect on p_f as seen in the figure. Although we see that P_{f_2} has a high probability in a packet based system, but as we will see later, the impact of P_{f_2} on system throughput will be scaled down by $2T_s/T$, which the result would be negligible.

Probability of collision and collision duration are other important metrics of the system. Fig. 4.9 depicts the probability of collision of a SU frame with different durations for our proposed scheme and for the conventional LBT scheme in a full duplex CR scenario. As it is seen, when the frame length increases, it is more likely that primary signal reappears within a SU frame interval, and that collision occurs. This collision probability is nearly the same for various false alarm probabilities (P_{f_2}) in the FDTR mode since the effect of P_{f_2} is scaled down by T_s/T factor according to (4.7). But in the LBT scheme, false alarm probability (P_{f_1}) has a considerable effect on the collision probability as seen in (4.8). In the conventional LBT scheme the probability of collision

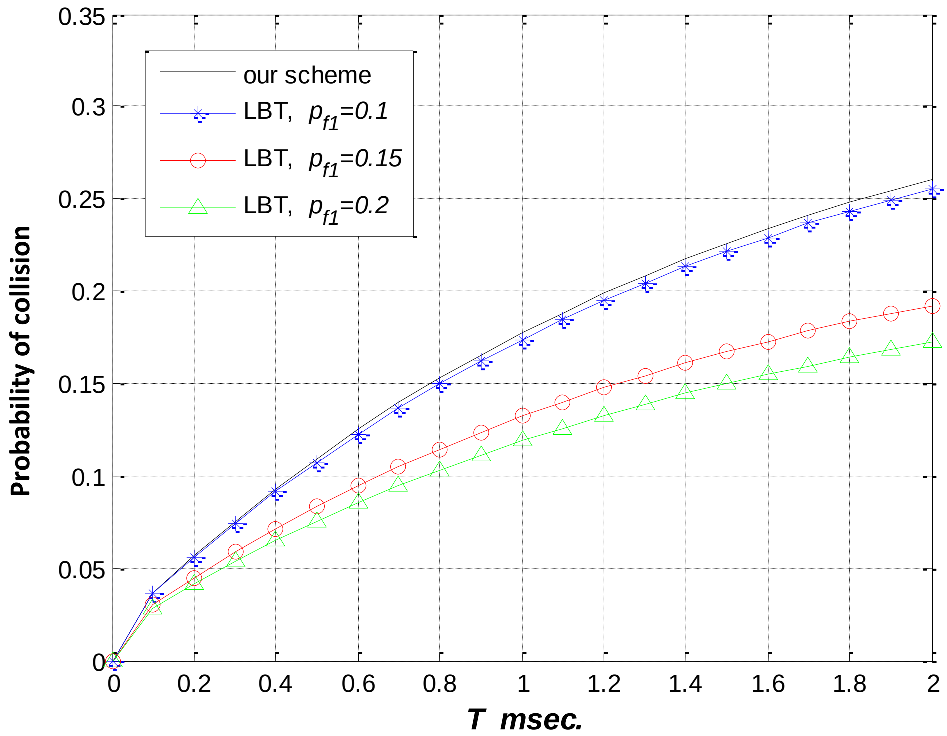


Fig. 4.9. Probability of collision vs. SU frame duration T for our scheme and LBT scheme with different values of P_{f_1}.

decreases as false alarm probability increases. However, this is not an advantage, since higher false alarm probability means more transmission opportunities lost and ends in network throughput decrease, which is not desirable.

Fig. 4.10 shows the average collision duration against SU frame duration for the synchronous and asynchronous schemes under different fading conditions of the PU-SU channels. When there is no fading, we have minimum average collision duration for the asynchronous mode of transmission (which is half of it for the synchronous mode). When we have fading, depending on the maximum Doppler frequency (f_m) of the channel, the average collision duration varies. For channels with high f_m, i.e. higher values for τ_m, outage probability is small and the average collision duration is near to no fade case. But for lower values of f_m (i.e. small values of τ_m), the possibility of deep fade durations increases, and this will result in higher average collision durations, as is depicted in this figure.

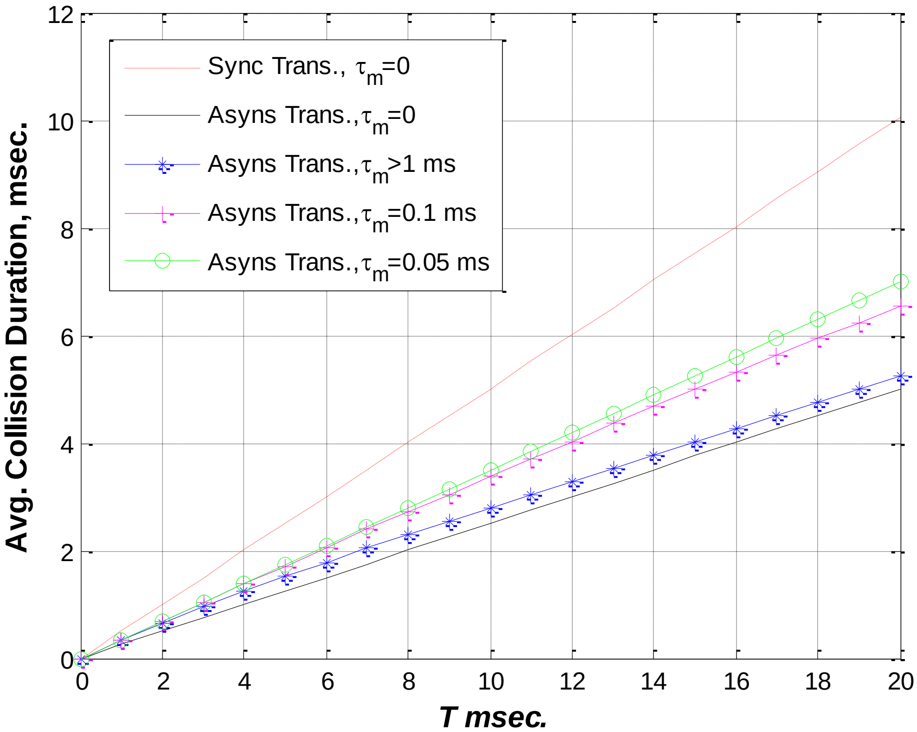


Fig. 4.10. Average collision duration vs. SU frame duration T for asynchronous and synchronous modes.

In Fig. 4.11, we have shown the normalized average cumulative collision duration. Again we see that false alarm probability has almost no effect on this metric in our proposed asynchronous method, but affects the same in the LBT scheme. We experience lower average collision probability when we check the presence of PU before every SU transmission in the conventional LBT schemes (which was shown in Fig. 4.5). However, as it is obvious from this figure, the average cumulative collision duration (which is more important in meeting PUs' protection criteria) in our proposed scheme is less than that in the LBT method.

Another important performance metric in a CR network is throughput. Fig. 4.12 shows a comparison between the achievable total throughput of the secondary network (throughput of both nodes) in the asynchronous and synchronous modes of operation at different values of frame duration. We observe that the simulated average throughput is very close to the theoretical one derived in (4.20) which verifies the validity of analysis.

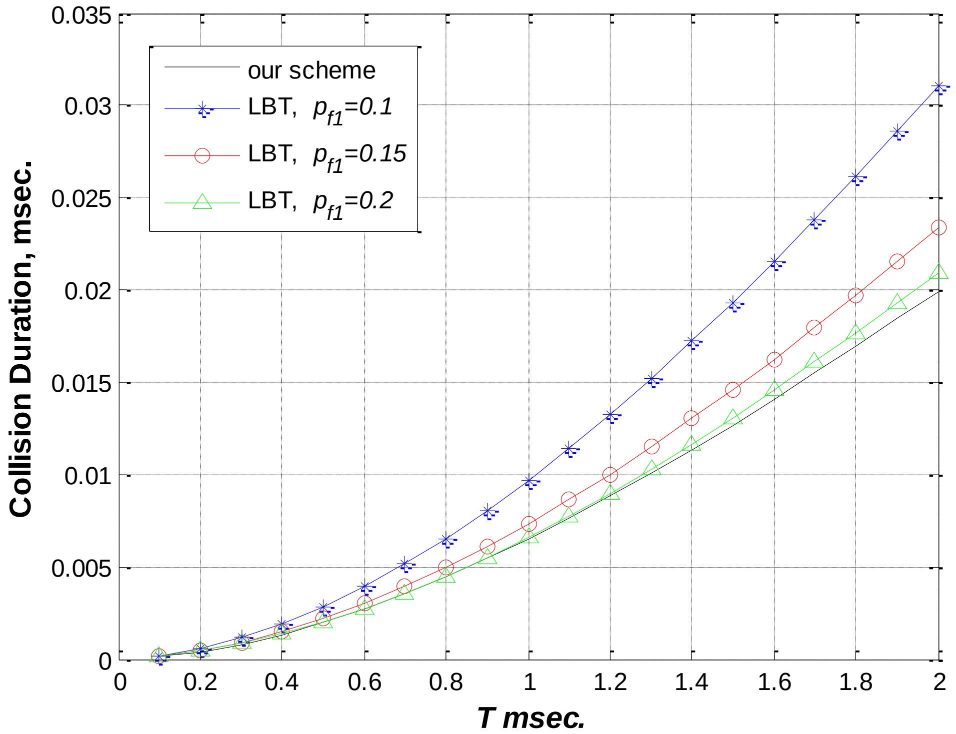


Fig. 4.11. Average cumulative collision duration vs. frame time T for proposed scheme and LBT scheme with different p_f.

We also observe that this average in the asynchronous mode is slightly more than in the synchronous transmission which is due to the shorter collision duration in the asynchronous mode. Moreover, we observe the effect of SIS factor β on throughput; poorer SIS capability (higher β) results in throughput degradation which is the expected result.

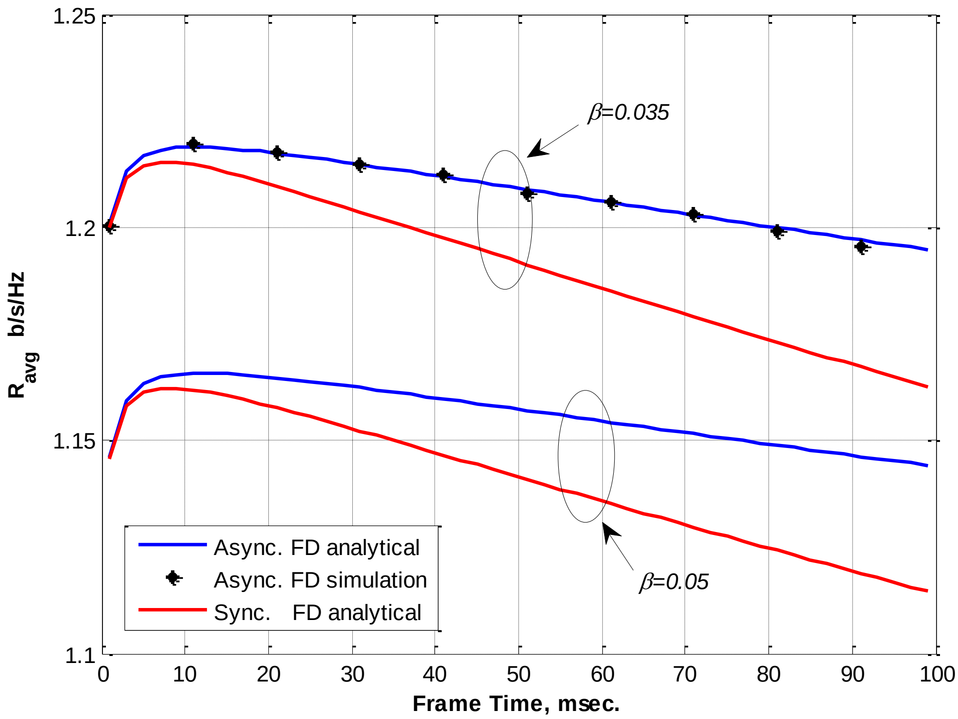


Fig. 4.12. Normalized average throughput for asynchronous and synchronous modes.

In Fig. 4.13, we see the variation of the secondary network throughput with SNR of the secondary network for the asynchronous and synchronous modes of our scheme. For comparison, the throughput of the same system with LBT scheme is plotted in this figure as well. It is evident that our protocol outperforms the LBT scheme by far, and the asynchronous mode of operation is slightly better than the synchronous mode at all values of secondary network SNR.

And finally Fig. 4.14 shows the SU network throughput in our proposed scheme and the LBT scheme, vs. SIS factor β. Here we see a close match between the analytical and simulation results. As it is seen, our method outperforms the fixed sensing method

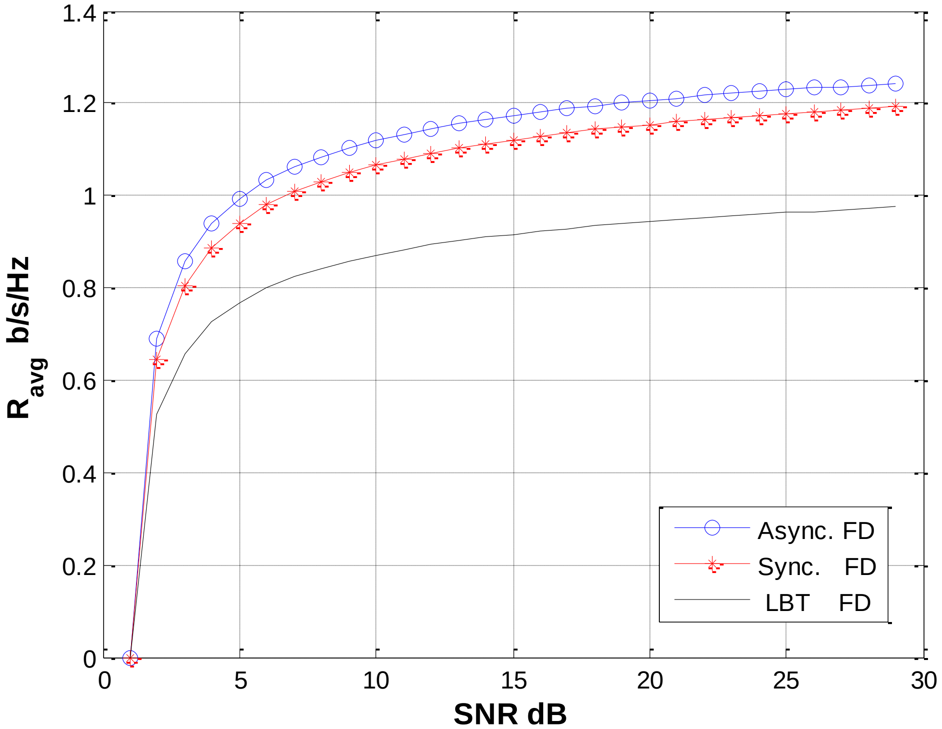


Fig. 4.13. Average SU throughput for our scheme and LBT scheme vs. secondary network SNR.

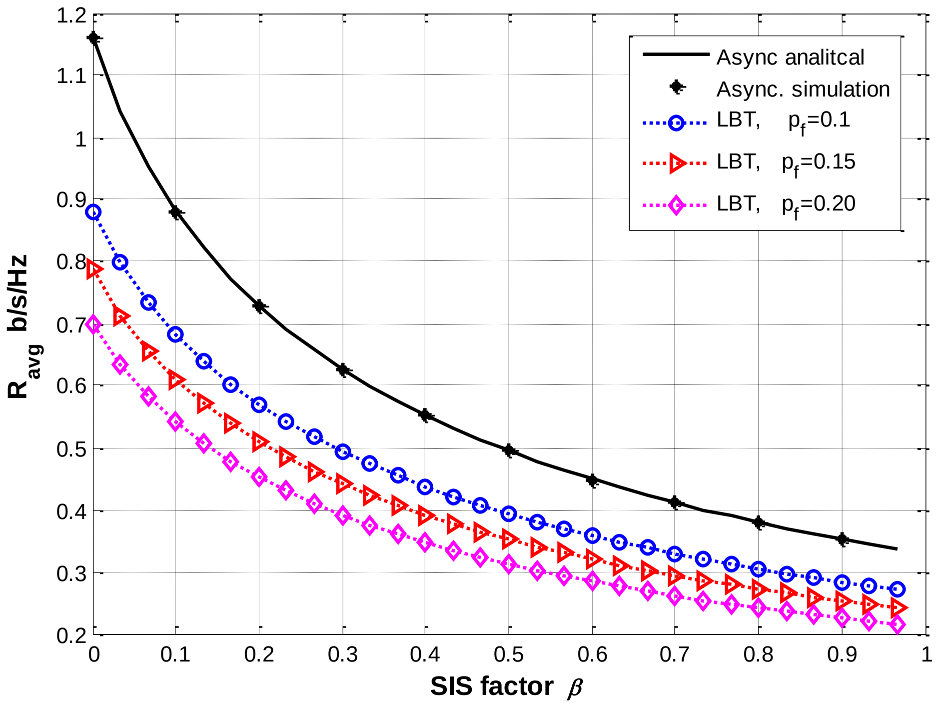


Fig. 4.14. Average SU throughput of our scheme and LBT scheme for different values of SIS factor β

decreases, the average throughput of the CR network monotonically decreases as well. In the LBT schemes, SIS factor, β, only influences the SU throughput through its effect on R_0 and R_1 according to (25), but in our protocol it has an additional effect on the false alarm probability, P_{f_2} according to (6). However, as shown in (20) and (21), the effect of P_{f_2} on the SU throughput in our protocol is negligible, which is in line with the behavior of SU throughput in our scheme and in the conventional LBT schemes

4.6 Summary

In this chapter we studied the full duplex paradigm in an opportunistic spectrum access cognitive radio network in which the secondary users are assumed to be capable of perfect or partial self-interference suppression. The secondary users in this paradigm operate in a bidirectional full duplex mode, which is less studies in cognitive radio systems.

The system consists of one pair of primary users whose activity is modeled by an alternating ON/OFF Markov model, and one pair of secondary users. We proposed a novel scheme incorporating cooperative sensing, and asynchronous full duplex communication carried out by secondary users. The proposed system comprises of two operation modes; the cooperative sensing (CS) mode when the primary network is active and the full duplex transmit and receive (FDTR) mode when PUs are not active. As a result of cooperative sensing in the CS mode, the probability of detection is high enough to guarantee no collision with the primary signal when the primary network is in ON state. Detection in the FDTR mode is based on collision event, when a primary signal returns back to the channel. In the FDTR mode, asynchronous transmission of the SUs results in decrease of probability of collision of primary and secondary networks, and similar decrease of the average collision duration, compared to that in synchronous systems, or traditional LBT schemes. We derived analytical closed form expressions for the probability of collision, the probability density function of collision duration, and the average throughput of the secondary network, for the proposed scheme and traditional LBT system for comparison. Extensive numerical simulations in the latest section of this chapter indicated that the proposed scheme outperforms both the conventional LBT and the synchronous transmission scheme in the FDTR mode, in terms of average collision duration and average secondary throughput.

Chapter 5

Contribution to IEEE 802.22 Standard

5.1 Introduction

Standards are published documents that establish specifications and procedures designed to maximize the reliability of the materials, products, methods, and/or services people use every day. They are based on industrial, scientific and consumer experience for the technologies and ideas which have been in use for a while and need guidelines to facilitate interoperability between products in a multi-vendor, multi-network and multi-service environment. Standards are regularly reviewed and revised after their publication, to ensure they keep pace with new technologies.

The development of the IEEE 802.22 WRAN standard is aimed at using cognitive radio (CR) techniques to allow sharing of geographically unused spectrum allocated to the television broadcast service, on a non-interfering basis, to bring broadband access to hard-to-reach, low population density areas, typical of rural environments, and is therefore timely and has the potential for a wide applicability worldwide. It is the first worldwide effort to define a standardized air interface based on CR techniques for the opportunistic use of TV bands on a non-interfering basis. It was published in 2011 and since then, revisions and appendices have been made, and more revisions will be made based on contributions made by industry or researchers.

As a contribution to IEEE 802.22.3 standard, we offered a novel suggestion based on application of FD in a Cognitive Radio network. This meta contribution were submitted to IEEE 802.22 standard committee and respective working group, which were considered in its plenary meeting in Nov. 2016 and published by organizing committee accordingly [116].

In the following section we describe the essence of this idea without any mathematical analysis or numerical simulations. It goes without saying that this idea needs some changes and a lot of developments to become complete and mature, and the authors reserve the right to add, amend or withdraw material contained herein.

5.2 Use of Directional Antenna

Detection of active incumbent users in IEEE 802.22 standard is performed by all secondary network CPEs and Secondary BSs in a cooperative manner. The BS fuses the results of all CPEs' signal detection and makes the final decision about the channel occupancy. In case the BS decides that the channel is not available, it will not let any of the CPEs to communicate over the channel in question, and will switch to other available channels. It might happen that some CPEs are in the coverage area of an incumbent user but other CPEs are far from it. In such cases, some CPEs will detect an incumbent signal or beacon (in case of wireless microphones) and will inform the BS that the channel is not available. Hence, the cognitive BS will not allow transmission on the said channel to any CPE, although only one CPE has detected the primary's signal. This situation is depicted Fig 5.1.

As it is seen in Fig. 5.1, only a small area of the WRAN coverage area is affected by an incumbent signal. Hence, a solution which can exclude that area, and allows the secondary's transmission over the remaining areas would result in an enhanced secondary's performance and efficiency.

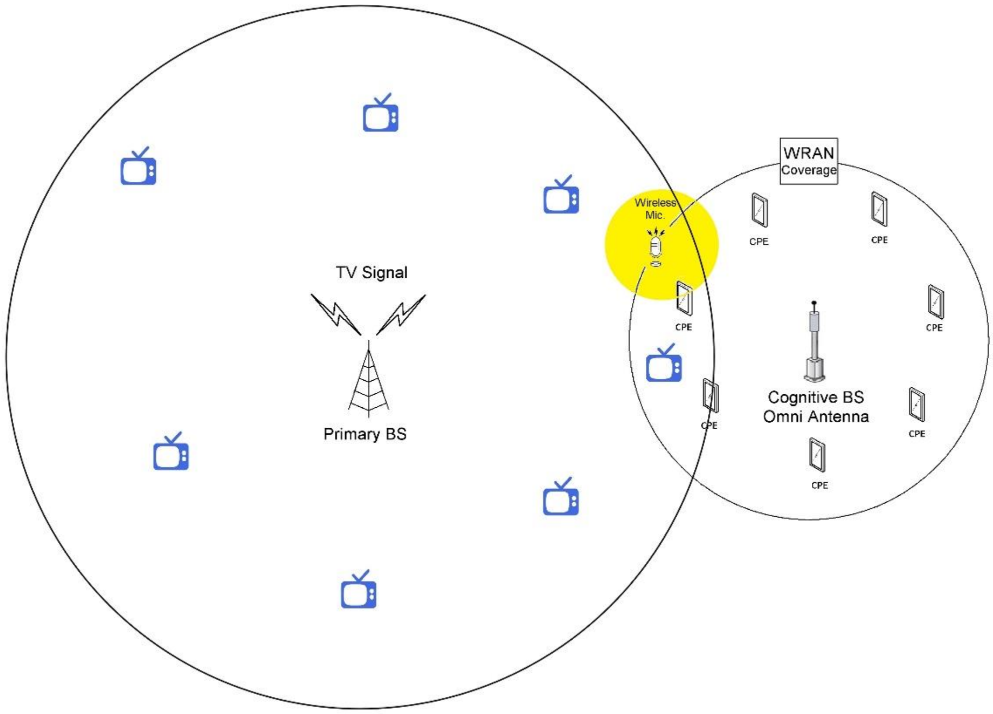


Fig. 5.1. Intersection of coverage areas of the primary and secondary networks.

If we equip the cognitive BS with directional sector antenna (such as those of WiMAX or mobile networks), the cognitive BS can differentiate between the CPEs experiencing interference, and those not in interference. Then it may limit its communication over this channel only with those CPEs which have not detected the incumbent signal, and are far from the coverage area of incumbent signal. Hence, the channel may be used for some CPEs and good throughput harvested without any interference. This scheme would specially be feasible when wireless microphones (with a small coverage area) are present. This scheme is depicted in Fig. 5.2.

In order to maintain fairness between all CPEs, or to meet QoS requirements for delay-constrained services such as voice or video, the cognitive BS may use another available channel for communication with those CPE's experiencing an incumbent signal in order to. This arrangement is shown in Fig. 5.3.

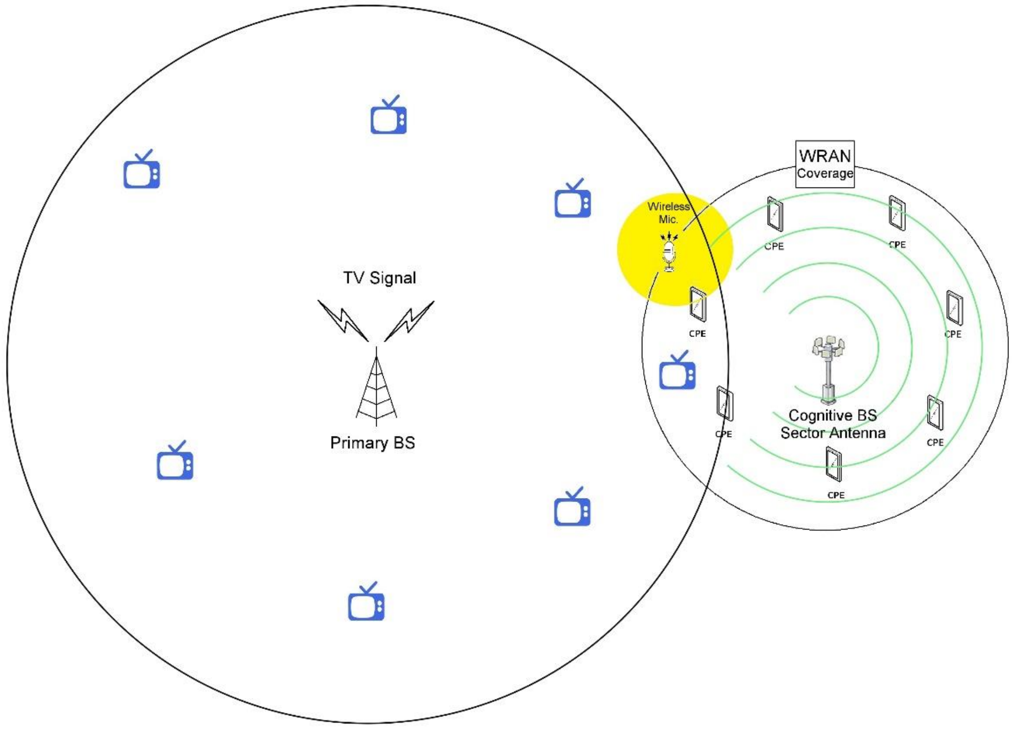


Fig. 5.2. Use of directional sector antennas in WRAN for selective sensing transmission.

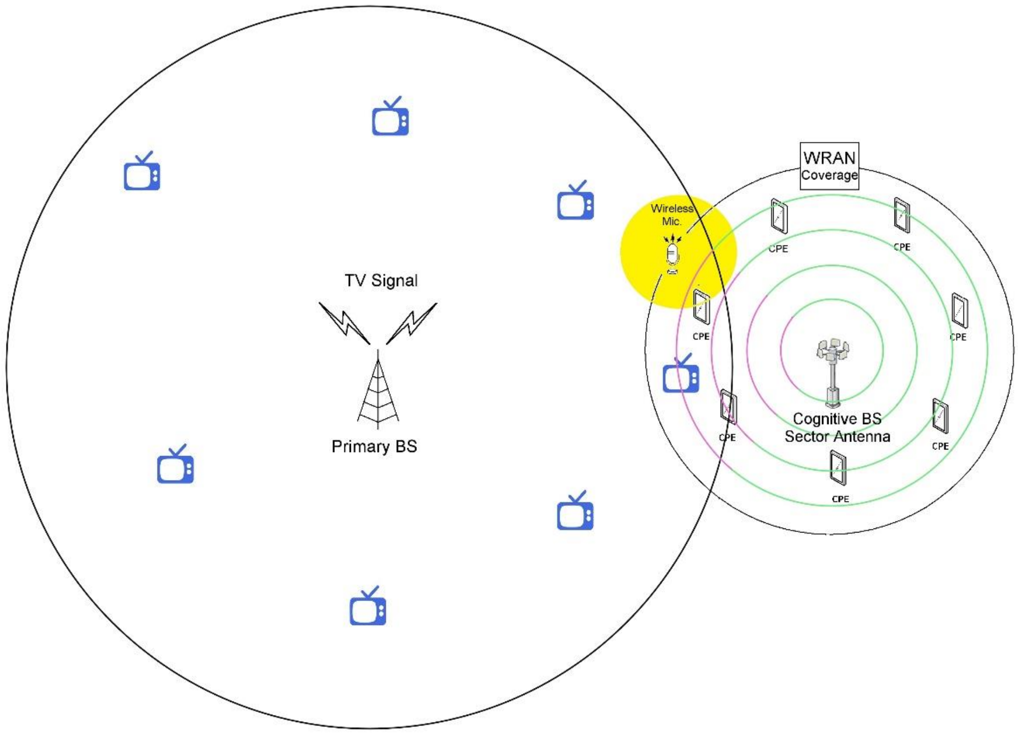


Fig. 5.3. Transmission over two different spectrums in a WRAN equipped with sector antennas.

This idea can be developed further to include FD operation of the cognitive BS which will increase the secondary network throughput and fairness of service. According to the published version of IEEE 802.22 standard, the BS and CPEs access the available spectrum in a time division multiple access (TDMA) order in a half duplex manner, similar to other HD centralized networks (e.g. WLAN). The BS grants access to the CPEs based on the priorities and algorithm defined in the MAC protocol. However, by equipping the BS with directional sector antennas, it can transmit and receive concurrently over the same spectrum in two opposite directions. This situation in shown in Fig. 5.4.

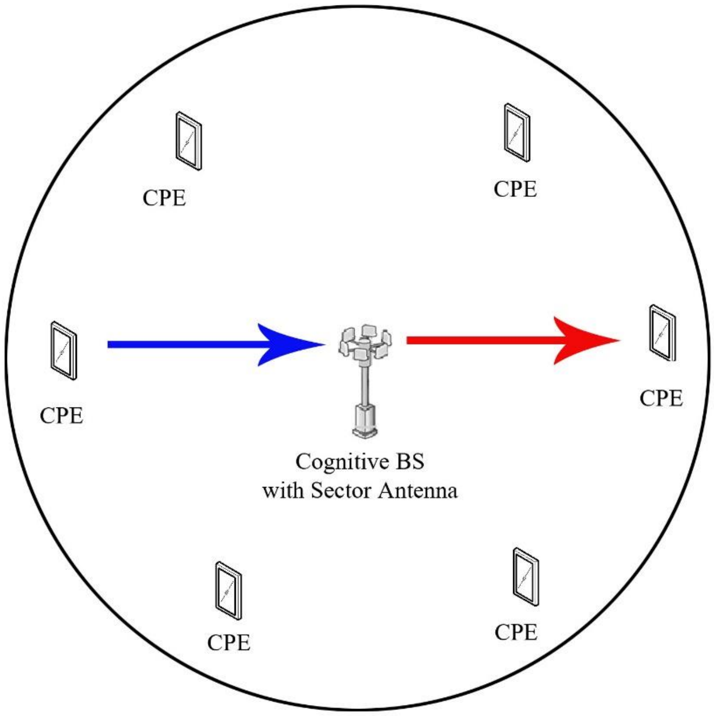


Fig. 5.4. FD operation of the cognitive BS equipped with sector antenna.

In this scheme, the BS should be capable of self-interference cancellation, but the CPEs do not need such capability. The interference from transmitting CPE on the receiving CPE is assumed to be low compared to the powerful transmission from BS to the CPEs, and may be considered as a noise and ignored. However, the BS has the capability to switch to HD mode whenever this interference is affecting the reception of the receiving CPE.

It might be possible that the BS operates in FD mode, even if the transmitting and receiving CPEs are not in opposite directions, as long as the interference of the transmitting CPE on receiving CPE is within acceptable range.

5.3 Future work with relevance to 802.22.3

Further work and consideration is needed with regard to the abovementioned 802.22 suggestion. However, various implications for 802.22.3 of such sectorised antennas, in the scope of sensing, can be derived. First, 802.22.3 should support the ability to define directional (in azimuthal plane) sensing and antennas thereof. It is suggested that there should be two modes of support specified:

- Fixed directional antenna sensing,
- Configurable/flexible directional antenna sensing.

Our reading of the current first pass submissions is that directional antennas are not currently supported (or at least not indicated as such), however, it would be good to add this capability.

The optimum number of sectors based on regional active incumbents and geographical databases is another parameter which should be investigated. As well, the list of inclination angles of the antenna in each sector, based on the location of CPEs in different altitudes should be prepared. As the cognitive BS operates in FD mode, its self-interference suppression capability and the acceptable range of suppression for proper operation of WRAN network should be studied in detail and specified for respective standard. The diagonal distance between opposite antennas at the BS would be considered in this connection.

Switching policy from FD to HD mode at the cognitive BS is another issue which would incorporates some optimizations in finding the optimum switching thresholds, based on the amount of interference impinged by transmitting CPE on receiving CPE, and BER or PER requirements of the services delivered in WRAN.

5.4 Summary

In this chapter we proposed the suggestion of use of directional sector antennas as a contribution to IEEE 802.22 standard, which may be considered in the future revisions and publication of the said standard, upon further investigations.

This suggestion which was accepted and published by corresponding standard committee as a primary meta contribution, shows how the ideas proposed in previous chapters of this thesis may be applied in practice.

The proposed idea needs a lot of work to become mature enough to be included in the standard eventually, which would be done as a part of our future work.

Chapter 6

Conclusion and Future Work

5.1 Concluding Remarks

Cognitive radio and full duplex communication are two promising technologies for enhancing spectrum utilization and combination thereof will improve the performance even further. This improvement will not only result in spectral efficiency, but would affect other metrics of a network positively. In fact, FD is a capability and CR is a networking paradigm. FD communications may replace almost all HD communications in all wireless systems for a better spectral efficiency, but CR idea may not be employed in all wireless platforms. Therefore, whenever we talk of full duplex cognitive radio, we mean a CR system in which FD communication is employed in some ways.

There are many advantages inherent to FD transmissions (as indicated in chapter 2) which if properly exploited will end in significant achievements in a network. The prime objective and contribution of this thesis has been the study of implementation of FD communication in CR networks and the consequent advantages and enhancements achieved therein. To accomplish this a brief study of the FD and CR technologies, their definitions, main functions, advantages and limitations, and their applications and complications was given in Chapter 2, followed by a literature review of the recent works on full duplex cognitive radios.

In Chapter 3 we considered an overlay CR network in which primary users where ARQ-enabled, and secondary network took the advantage of the opportunities arising during ARQ rounds of primary operation to transmit some data. There are some studies which have considered similar situation but all in half duplex mode, and were reviewed briefly in the chapter. Through proper application of FD transmissions in the said system, we proposed a cooperative ARQ scheme which not only inflicted no additional interference on primary's signals, but could enhance its performance in terms of packet error rate and throughput in the case of HARQ. In contrast to other cooperative schemes, primary users are oblivious in our proposed model and no coordination between primary and secondary networks is required. Within the course of analyses, we considered packet error rate as a practical and usable metric, and derived exact closed forms for average packet error rates of primary and secondary network in the proposed cooperative scheme, as well as respective throughputs for both full duplex ARQ and HARQ modes, as well as non-cooperative half duplex ARQ mode, for comparison purposes. Numerical results based on extensive simulations incorporating coded and uncoded modulation schemes verified the validity of the analyses, and demonstrated that the proposed scheme outperforms non-cooperative half duplex models studied before. The proposed scheme required less power (less than half) compared to other similar systems, to achieve a target PER or throughput

Chapter 4 investigated another design paradigm in an interweave cognitive radio network. Therein, a new scheme for secondary users'(SUs') sensing and transmission was proposed which comprised of two modes of operation; Cooperative Sensing (CS) mode and in-band Full Duplex Transmit and Receive (FDTR) mode. Cooperative sensing was applied to reduce misdetection probability and to avoid collision with primary users (PUs) during the primary network BUSY times. Upon successful detection of a spectrum hole through a cooperative MAC protocol, both transmitting and receiving SUs start to simultaneously transmit and receive date in FDTR mode. The secondary users which are capable of partial or perfect self-interference suppression (SIS), would communicate bidirectionally in an asynchronous FD manner in order to decrease average collision duration and average cumulative collision duration, compared to conventional LBT schemes in half duplex or full duplex modes. Analytical closed forms for probability of collision, average collision duration and cumulative collision duration, as well as throughput of SU network were derived and performance of the proposed protocol, its effectiveness, and advantages over conventional methods of sensing and transmission were analyzed and verified through numerical simulations. It was observed that the proposed scheme outperforms the conventional methods of sensing such as LBT in both HD and FD modes when an asynchronous transmission policy is employed.

In Chapter 5 a novel suggestion on the application of directional sector antennas capable of FD operation in a wireless regional area network was proposed which has been submitted to IEEE 802.22 standard committee as a contribution. It was shown that the proposed idea could result in higher spectral efficiency of the CRN, and providing fairness and QoS for services offered through WRAN.

6.2 Future Work

In this thesis we proposed two novel schemes on FDCR technology. However, there are still many topics to be explored and studied. As mentioned in previous chapters, FDCR is a young topic with huge research opportunities. Full Duplex transmissions may be applied in almost all half duplex CR design paradigms and significant results may be achieved.

On the other hand, although we have already shown considerable advantages achievable in the proposed schemes compared to half duplex and traditional models, there are many suggestions which may improve the performance of the proposed schemes in this thesis even further. In the proposed FD cooperative ARQ system in Chapter 3, we may implement network coding in order to achieve better rates for primary and secondary. Another suggestion would be cooperation of a non-oblivious primary network with the secondary users during the OFF cycles of primary network which is expected to result in higher throughput for secondary. Employing other types of Hybrid ARQ, such as Chase combining and type I incorporating FEC codes added to the packets may be considered in future works. Another suggestion would be combining the method proposed in [75] with our method to achieve further throughput for the secondary network while observing an interface budget.

Multi-user platforms have always been of great interest in wireless systems, which may be considered in both designs proposed in Chapters 3 and 4. For the system of Chapter 3, we may consider more than a pair of secondary users, and based on a MAC protocol allow the pair with best channel conditions or with higher priorities to take part in cooperation and taking advantage of ARQ round opportunities for their own benefit.

In the second proposed scheme, we have considered a single bidirectional channel for a pair of secondary users. This scenario may be extended to multi user and multi-channel case, where more than a pair of secondary users contend for accessing the channels, and the respective MAC protocol for channel access may be studied. Again the channel conditions or QoS priorities may be the criteria for selecting the winner pair. Switching policies to other channels and analysis of the network performance under such conditions are other suggestions for a multi-channel platform which could be investigated in future works. In addition, we may add a transition mode between the CS and FDTR modes in which the communication between SUs would be one way (half duplex), but the source node will sense the channel in a FD manner at the same time. This mode will add to the protection of the primary network whenever the interference from primary's transmission on the secondary's communication in the FDTR mode is not enough to trigger a NACK signal and forces the CRN to stop transmission.

As the Chapter 5 is a primary suggestion for a possible enhancement in operation of a WRAN, there are a lot of work to be done in order to make the idea mature and applicable. First of all, an extensive study on the parameters of the proposed sector antenna for the secondary BS, such as the gain of antennas, number of antennas, the optimum inclination and azimuth angles is required. As the cognitive BS is assumed to operate in FD mode, the acceptable range of SIS factor for the BS and suitable method for self-interference cancellation should be specified in future development of this idea.

Respective MAC layer algorithms for directional sensing and transmission, selection of the transmitting and receiving CPEs in the FD mode, and switching policies from HD to FD modes of operation at the BS are other important research topics which could be investigated in future.

Appendix

Rayleigh Flat Fading Channel Modeled as a $G/H_R/1$ Queue

A.1 Introduction

Design and implementation of an efficient and practical transmission over wireless fading channels requires accurate and predictable channel models. But the existing physical layer channel models (e.g. Rayleigh fading model with a specified Doppler spectrum) do not explicitly characterize a wireless channel in terms of delay, delay violation probability and minimum sustainable data rate. As most of the practical communication systems are delay constrained (e.g. voice and video content), we need a cross layer model that interprets delay and delay violation probability (MAC layer metrics) in terms of time varying channel conditions (physical layer characteristics of the channel).

The need to quantify the performance of delay-constrained systems gave rise to the notions of *outage* and ϵ-*capacity* [118]. An outage event occurs when the attempted transmission rate is greater than the *instantaneous capacity* of the channel which is the capacity for a particular set of fading states affecting a codeword. Outage results in decoding errors at the receiver. ϵ-capacity is the maximum rate that can be sustained with the probability of outage no greater than ϵ. But these notions do not consider MAC layer delay in an explicit way. Cross layer schemes incorporating Physical and MAC layers have shown improvements in QoS provisioning [121], [122] but such methods should be formulated for each specific problem. Wu and Neggi in [119] offered a new model which combines physical and MAC layers into a single layer (link layer) and proposed the notion of effective capacity which is the dual of effective bandwidth in wired systems. Their model is a useful tool which considers delay and delay violation probability in terms of a constant maximum input rate to a wireless system. But their model incorporates an algorithm for estimation of channel parameters and has no analytic closed form expect for a special case of low signal- to-noise ratio.

On the other hand, queuing theory provides us with exact and well-defined relations between arrival process, service process and waiting time in a service facility queue [120]. In general, we may consider any transmission system as a *G/G/m* queuing system where relations between waiting time, service time, busy period, and characteristics of input traffic and transmission facility in a general form are well formulated. However, in this approach, it is difficult to formulate the arrival and service processes in an exact mathematical format except for a few special cases. Hence, subsequent complex procedures for obtaining service time and waiting time distributions are not straightforward and in many cases not possible.

In this appendix we have proposed a novel method for expressing the service characteristics of a Rayleigh flat fading channel based on parallel combination of some exponential distributions whose mathematical behavior are exactly known. This simple analytic cross layer queuing model enables us to derive QoS metrics of the system (i.e. delay, minimum sustainable rate, and delay violation probability) from waiting time distribution of the system which may be easily obtained for a channel with known marginal probability distribution and input process characteristics. To best of our knowledge no general analytic cross-layer model for a wireless channel has been proposed so far and this work can be regarded as a novel contribution to channel modeling. Although in this article we have considered a Rayleigh channel, this method can be easily extended to other channel types such as Ricean or Nakagami-m.

A.2 System Model

We consider a simple downlink transmission system in an uncorrelated Rayleigh flat fading channel. In this fluid model, arrivals and departures to and from system are continuous and data units are bits or nats. Once we have found the exact model for a single channel, then it can be easily extended to multi-channel or multi-user systems. In a continuous system, transmission of a data unit occurs within dt seconds. In discrete systems, dt is replaced by a transmission block with a duration of T seconds and channel status remains the same during a transmission block (assuming a flat fading channel type), but may vary from block to block independently. We assume that transmitter has the capability to dynamically adapt its rate to be close to instantaneous Shannon capacity of the channel at every transmission instant.

We model this system with a *G/G/1* queue as shown in Fig. A.1. The input traffic has an arbitrary interarrival time distribution $a(t)$. The service times are also independently drawn from an arbitrary distribution given by $b(t)$. There is one server (transmitter) available. Data units (bits, nats or packets) are queued in a buffer prior to transmission and service is offered in a First In First Out (FIFO) order. Once a data unit is on-air by the transmitter it is regarded as being served. As we presumed transmitter is always transmitting at channel instantaneous rate, then it is logical to consider all data units are received at the receiver without error and no retransmission occurs.

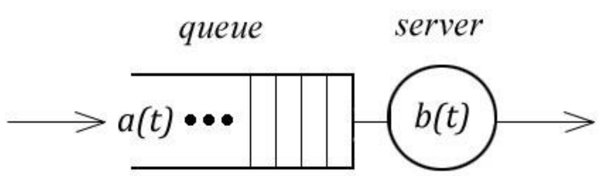


Fig. A.1. System model as a G/G/1

Whenever we face a queue in the system, we need to know how long on average a data unit waits in the queue before being served, and what is the probability of waiting

t seconds in the queue i.e. waiting time distribution $w(t)$. Of course $w(t)$ is dependent on arrival distribution; $a(t)$, and departure time distribution; $b(t)$.

Input traffic (arrival process) in communication systems is usually modeled by some known processes such as Poisson, Pareto, or a self similar distribution. In our model the service process $b(t)$ follows the instantaneous capacity of wireless channel itself, which has a random behavior and we have assumed that coding and modulation schemes are so flexible to follow it dynamically.[1] Therefore $b(t)$ may be regarded as a random function of wireless channel characteristics. Then our first step in modeling the channel would be to obtain $b(t)$ for a given channel.

As mentioned earlier, our channel is Raleigh fading, but as will be seen later, this method can be easily extended to other types of channels such a Ricean or general Nakagami channels. In terms of queuing theory notation, we consider a *G/Rayleigh/1* queue.

A.3 Service Time Distribution of G/Rayleigh/1 Queue

In an uncorrelated Rayleigh fading channel with additive white Guassian noise (AWGN), the channel gain stays fixed over each transmission block and may vary from block to block independently. Let $h(t)$ be the time varying channel gain and $h(t)^2 = x$ the corresponding power gain with the following probability density functions respectively

$$p_H(h) = \frac{1}{\sigma^2} e^{\frac{-h^2}{2\sigma^2}} \quad , \quad p_X(x) = \frac{1}{2\sigma^2} e^{\frac{-x}{2\sigma^2}}. \tag{A.1}$$

Then according to Shannon formula instantaneous channel capacity would be

$$c(t) = Wln\left(1 + x(t)\frac{P_0}{WN_0}\right), \tag{A.2}$$

[1] In practice this assumption is relaxed by considering a discrete version of the channel with a limited number of transmission rate steps.

where $c(t)$ is the instantaneous capacity of channel at t in nats. W is the channel bandwidth, P_0 is the fixed transmission power, and N_0 is the noise power. Without loss of generality we assume $W = 1$ and denote the average SNR as $\frac{P_0}{N_0} = \gamma$. Then we can rewrite (A.2) as $c(t) = ln(1 + \gamma x(t))$.

Service provided by the channel in t seconds would be

$$S(t) = \int_0^t c(t)dt. \tag{A.3}$$

Given the distribution function of x, we can find the pdf of transmission rate as

$$p_{C_t}(c) = \beta e^{(c_t + \beta - \beta e^{c_t})}, \tag{A.4}$$

where $\beta = \frac{N_0}{P_0 2\sigma^2}$ .The cumulative distribution function (cdf) of (4) is

$$P_{C_t}(c) = 1 - e^{\beta - \beta e^{c_t}}. \quad \text{A.5}$$

Equation (A.4) is depicted in Fig. A.2 for different values of β. β is influenced by fading status of the channel. When $\beta = 0$ channel has no fading and acts like a

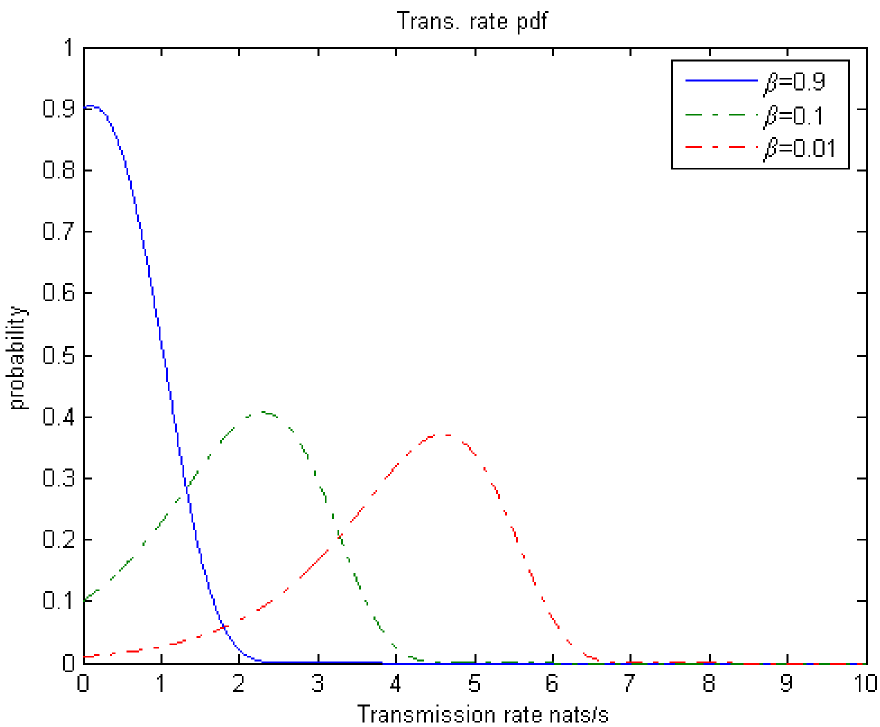


Fig. A.2. Transmission rate pdf in a Rayleigh channel deterministic channel. While β increases, the average SNR of the channel drops and the channel experiences increasing levels of fading. When $\beta = 1$ the channel is in deep fade and no transmission takes place.

In a discrete system with transmission blocks of duration *T*, the amount of service provided by the system through the channel in *n*th transmission block (s_n), and cumulative service in *k* successive blocks (S_k) would be:

$$s_n = c_n T \,, \qquad S_k = \sum_{i=1}^{k} s_i. \tag{A.6}$$

where c_n is the discrete probability distribution of the channel capacity. In (A.4) and (A.5) subscript *t* should be replaced by *n* to get the discrete form of pdf and cdf of channel capacity respectively.

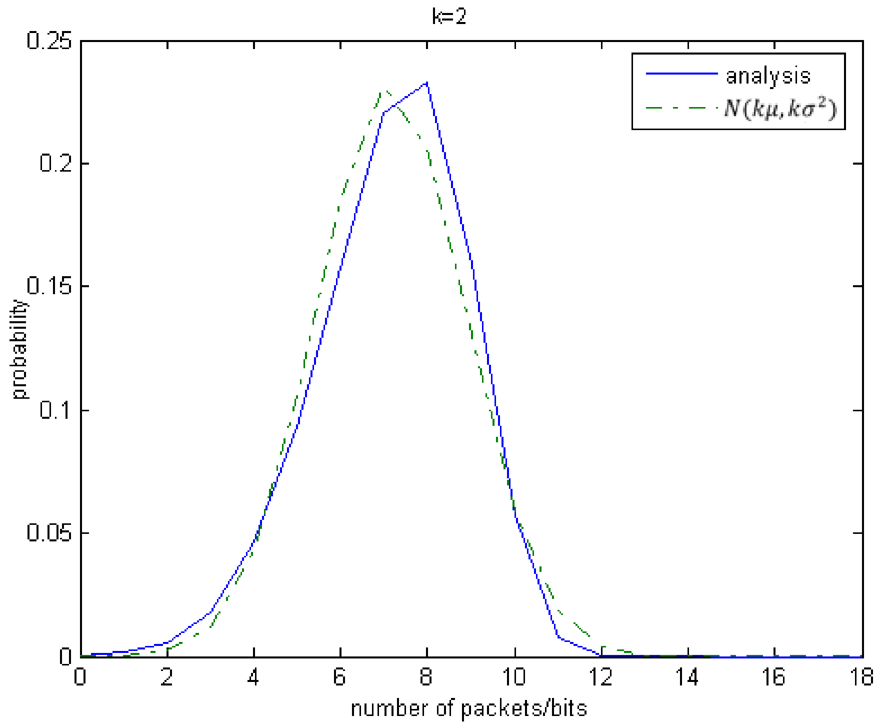
Probability density function of the amount of service in k successive blocks in discrete form $p_{S_k}(S)$ would be the *k*-fold convolution of the channel capacity pdf (A.4) which is:

$$p_{S_k}(S) = p_{c_n}(c) * p_{c_n}(c) * \dots . p_{c_n}(c). \tag{A.7}$$

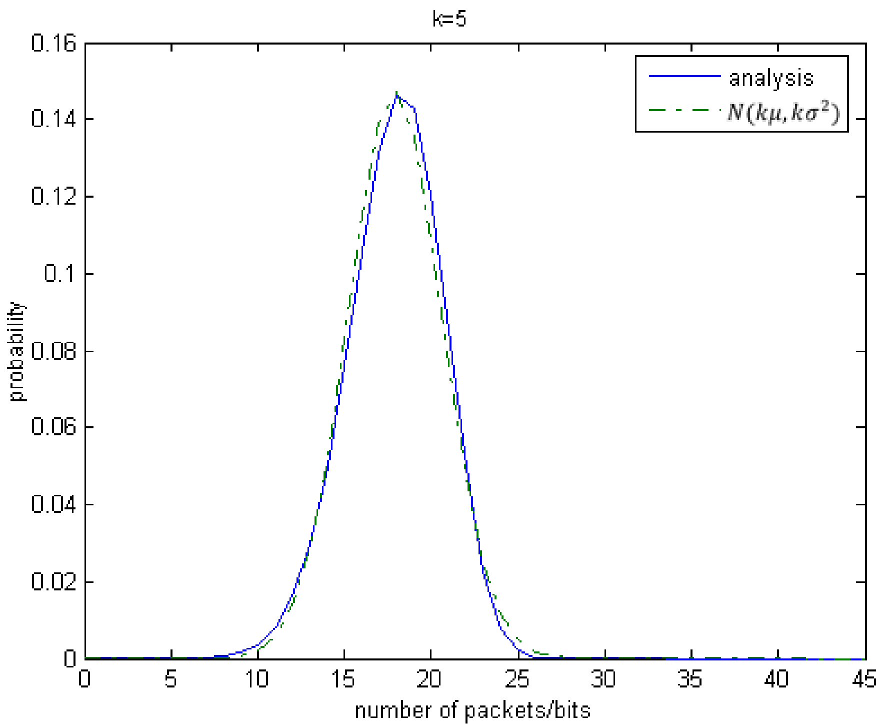
In general, it is not possible to calculate (A.7), but it can be replaced with a close approximation. In Fig. A.3 we have plotted this probability density function for k=2 and k=5. As it was predictable, according to the Central Limit Theorem (CLT), this probability density function will converge to a normal distribution with mean $k\mu$, and variance $k\sigma^2$ (i.e. $\mathcal{N}(k\mu, k\sigma^2)$)which is given in (A.8).

$$p_N(n) = \frac{1}{\sqrt{2\pi k}\sigma} e^{\frac{-(n-\mu k)^2}{2k\sigma^2}}. \tag{A.8}$$

On the other hand, μ and σ^2 are respectively the mean and variance of (A.4) which is transmission rate distribution for a Rayleigh channel which could be found by numerical methods.



(a)



(b)

Fig. A.3. Probability of number of transmitted packets in *k* blocks; (a) *k*=2, (b) *k*=5.

As mentioned before we need service time distribution of the system. Instantaneous capacity of channel indicates number of data units served in each second, so we can regard $1/c(t)$ as the service time of each data unit. Having the pdf of $c(t)$ in (5), then we can find pdf of $t = 1/c(t)$ as

$$b(t) = \frac{\beta e^{(1/t+\beta-\beta e^{1/t})}}{t^2}. \qquad \text{(A.9)}$$

Equation (A.9) has been depicted in Fig. A.4 for a Rayleigh channel with $\beta = 0.5$ and and $\beta = 0.1$.

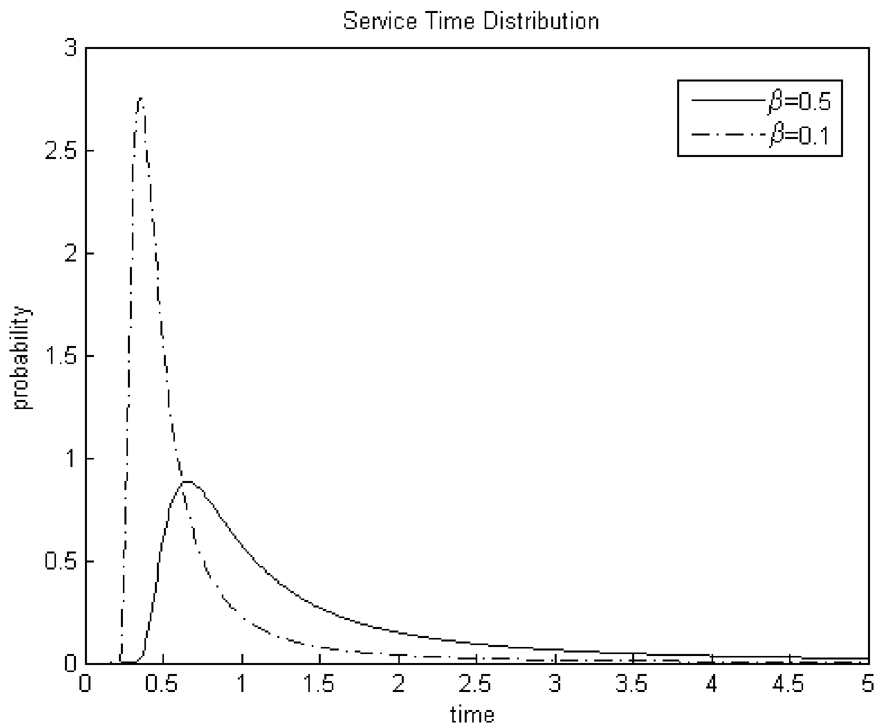


Fig. A.4. Service time distribution

A.4 $G/H_R/1$ Queue Model

In previous section we obtained the service time probability distribution function of a perfect transmitter in a Rayleigh fading channel. Traditionally, in order to find the fundamental metrics of a queue with a general service process, we need to derive its Laplace transform and if its intact form exits, subsequently find the Laplace transform of waiting time distribution through complicated procedures which are applicable just to special cases [120]. As it is obvious there is no closed form for Laplace transform of (A.9) so we cannot continue the classical method. But there is another approach for analyzing a *G/G/1* system.

In general, any queue with a given service process can be decomposed into a combination of series exponential processes in parallel [120]. A simpler case is the combination of R parallel exponential processes as in figure A.5.

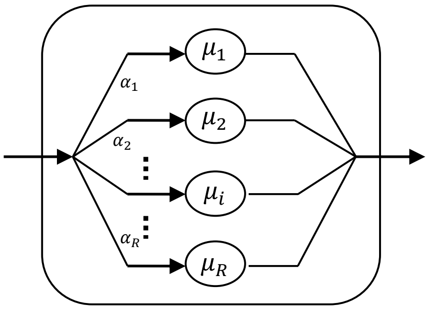


Fig. A.5. Decomposition of a general service facility

The service time probability corresponding to the combination of R parallel exponentials as shown in Fig. A.5 can be written as

$$b(t) = \sum_{i=1}^{R} \alpha_i \mu_i e^{-\mu_i t} \quad ,t \geq 0 \quad and \quad \sum_{i=1}^{R} \alpha_i = 1 \tag{A.10}$$

In this form we can easily find the Laplace transform of service facility as:

$$B^*(s) = \sum_{i=1}^{R} \alpha_i \frac{\mu_i}{s+\mu_i}. \tag{A.11}$$

The pdf given in (A.10) is referred to as *hyperexponential* distribution and is denoted by H_R. Theoretically any continuous pdf on $[0,\infty)$ which decays exponentially can be uniformly approximated on $[0,\infty)$ by a generalized hyperexponential pdf [122]. According to (A.9) and as it is seen in Fig. A.4, the service time distribution of a Rayleigh channel is an exponentially decaying process, and accordingly, it can be decomposed into a superposition of R exponentials. In order to employ this theorem, we need to break down (A.9) in the form of (A.10).

For this purpose, we use the so called *Prony's* method [124]. Similar to the Fourier transform, Prony's method extracts valuable information from a uniformly sampled signal and builds a series of damped complex exponentials or sinusoids. Let $f(t)$ be a signal consisting of *N* evenly spaced samples. Prony's method fits a function

$$\hat{f}(t) = \sum_{i=1}^{N} \frac{1}{2} A_i e^{\mp j\phi_i} e^{\lambda_i t} \tag{A.12}$$

to $f(t)$ where $\lambda_i = \sigma_i \pm j\omega_i$ are the eigenvalues of the system, σ_i are the damping components, ϕ_i are the phase components, $A_i s$ are the amplitude components of the series, and j is the imaginary unit $(j^2 = -1)$. Prony's method is essentially a decomposition of a signal with *N* complex exponentials via a numerical process.

This decomposition in general involves complex exponentials. The argument is that whereas this corresponds to no physically realizable exponential stage, so long as their poles are in complex conjugate pairs then the entire service facility will have a real pdf, which corresponds to the feasible cases. Usually a two or three stage hyperexponential function gives a close approximation to the main function. In the following we have considered a two stage approximation for simplicity, although this procedure can be easily extended to three stages as well.

Approximating (A.9) with a two-stage hyperexponential system, the resulting pdf is of the following form

$$\tau_1 e^{-\mu_1 t} + \tau_2 e^{-\mu_2 t}, \tag{A.13}$$

where coefficients (τ_1, μ_1) and (τ_2, μ_2) are found through numerical (e.g. Prony's) method, and we should transform it into $\alpha_1 \mu_1 e^{-\mu_1 t} + \alpha_2 \mu_2 e^{-\mu_2 t}$ where $\alpha_1 = \frac{\tau_1}{\mu_1}$ and $\alpha_2 = \frac{\tau_2}{\mu_2}$. Since (A.13) is the decomposition of a pdf, then its integral over $(0, \infty)$ would be 1. The same is true about $\mu_1 e^{-\mu_1 t}$ and $\mu_2 e^{-\mu_2 t}$ which are exponential pdfs. Therefore $\alpha_1 + \alpha_2$ will be equal to one, which is the required condition of (A.11).

The last step would be finding the waiting time distribution of the system. The waiting time distribution depends on arrival process so we require knowledge of the arrival process. Here we consider the simplest case which is a Poisson arrival.

According to P-K transform equation, the Laplace transform of waiting time distribution in a M/G/1 queue is [120]:

$$W(s) = \frac{s(1-\rho)}{s-\lambda+\lambda B(s)}. \qquad \text{(A.13)}$$

Inserting (A.11) in the above formula we get

$$W(s) = \frac{s(1-\rho)}{s-\lambda+\lambda\left(\alpha_1 \frac{\mu_1}{s+\mu_1} + \alpha_2 \frac{\mu_2}{s+\mu_2}\right)}, \qquad \text{(A.14)}$$

and after some modifications it will be

$$W(s) = \frac{s(1-\rho)}{s-\lambda+\lambda\left(\alpha_1 \frac{\mu_1}{s+\mu_1} + \alpha_2 \frac{\mu_2}{s+\mu_2}\right)},$$
$$W(s) = \frac{s(1-\rho)(s+\mu_1)(s+\mu_2)}{(s-\lambda)(s+\mu_1)(s+\mu_2)+\lambda\big(\alpha_1\mu_1(s+\mu_2)+\alpha_2\mu_2(s+\mu_1)\big)}. \qquad \text{(A.14)}$$

By reordering denominator of (A.14) in terms of s, and after some simplifications we will have

$$W(s) = \frac{(1-\rho)(s+\mu_1)(s+\mu_2)}{s^2+s(\mu_2+\mu_1-\lambda)+(\mu_1\mu_2-\lambda\mu_2-\lambda\mu_1+\lambda\alpha_1\mu_1+\lambda\alpha_2\mu_2)}. \qquad \text{(A.15)}$$

Dividing the numerator by denominator to reduce the degree of the numerator by one, and then carrying out partial-fraction expansion, we find the inverse Laplace transform of (A.15) which can be written as

$$w(t) = k_1 u_0(t) + k_2 e^{-\mu_1 t} + k_3 e^{-\mu_2 t} \qquad t \geq 0, \quad (12) \qquad \text{(A.16)}$$

where k_1, k_2, and k_3 are functions of α_1 and a_2.

A.5 Numerical Results

In this section we have considered a Rayleigh fading channel with sampling time of 10 microseconds and maximum Doppler frequency of 100 Hz. We have considered three scenarios of fading with average SNR of 1.85, 4 and 20 dB respectively. With the assumption of DPSK modulation, we have simulated the response time of a Rayleigh channel in the said scenarios, and applied Prony's method for approximating the channels with a second-order hyperexponential process.

The results are shown in Figs. A.6, A.7 and A.8 for severe, medium and weak fading channel conditions respectively.

As it is seen in these figures, the service time probability is well approximated by sum of two exponentials. In the first, second and third scenarios, the resulting exponentials are respectively:

$$\begin{aligned} b_1(t) &= 0.479\,(0.2296e^{-0.2296t}) + 0.535(0.557e^{-0.551t}), \\ b_2(t) &= 0.212\,(0.1233e^{-0.1233t}) + 0.803(1.040e^{-1.040t}), \\ b_3(t) &= 0.028\,(0.031e^{-0.031t}) + 0.938(2.500e^{-2.500t}). \end{aligned} \tag{A.17}$$

As can be seen, when channel condition improves and approaches the AWGN channel without fading, there is a dominant term with a very fast decay rate which is equivalent to a channel with a very short waiting time.

As it is seen in the figures, there is a close match between analytical and simulation results which verifies the model presented.

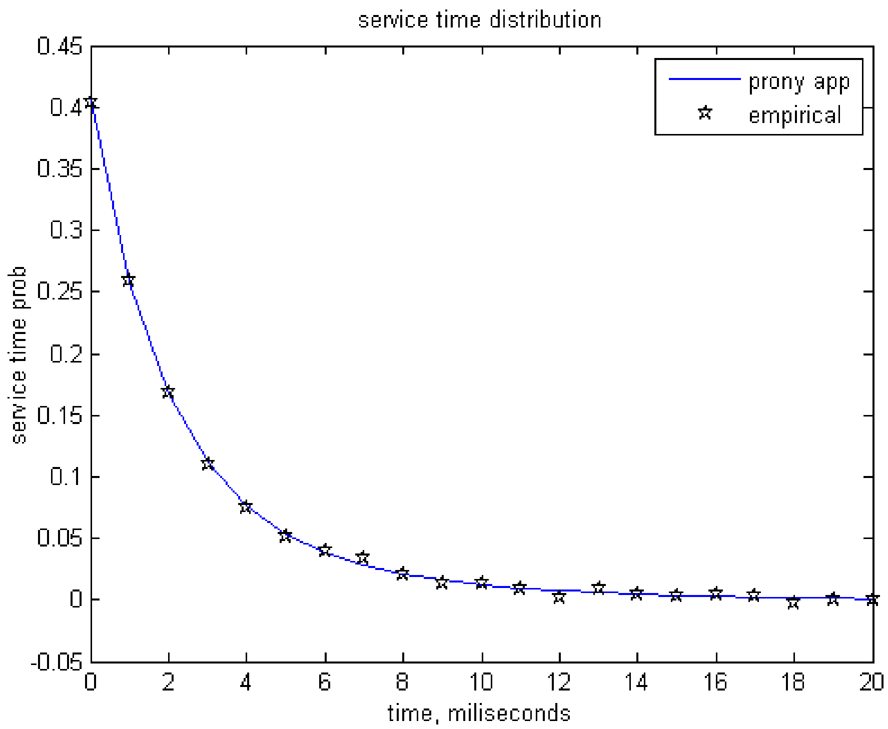


Fig. A.6. Service time distribution for SNR= 1.8 dB.

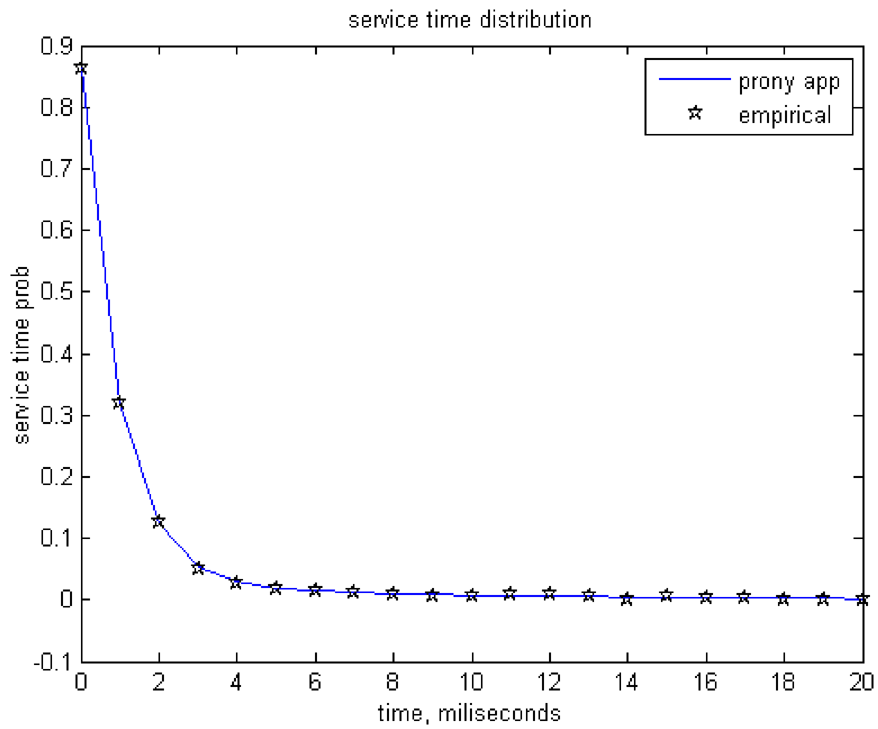


Fig. A.7. Service time distribution for SNR= 5 dB.

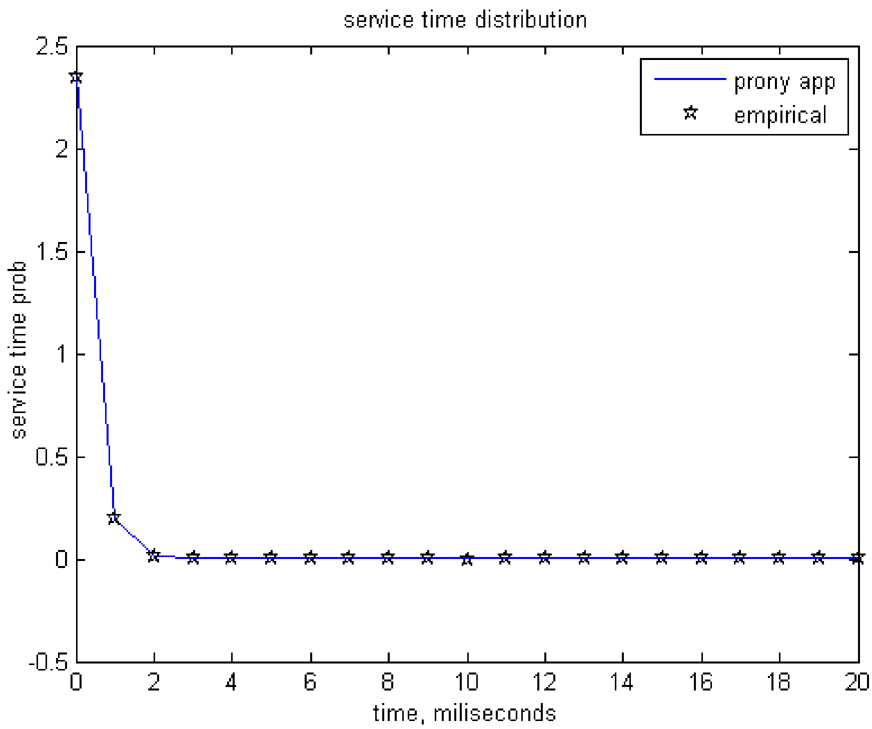


Fig. A.8. Service time distribution for SNR= 20 dB.

A.6 Conclusion

In this article we proposed a novel queuing model for a Rayleigh flat fading channel. First we found the service rate and accordingly the service time distribution of a transmitter with CSI, and then using Prony's method, modeled the system as a G/H_2/1 queue. This queue model enables us to derive QoS metric of the system based on waiting time distribution. Following numerical illustrations verified the validity and applicability of our proposed model with a close approximation. Solution developed in this paper, although derived for a Rayleigh fading channel, can be extended to other channel types as well.

References

[1] Qualcomm Corp." Evolution of wireless applications and services", white paper, Dec 2007.

[2] Cisco Corp." Global Mobile Data Traffic Forecast Update, 2015–2020 White Paper", Feb 2016, Document ID 1454457600809267.

[3] IEEE 802 LAN/MAN Standards Committee 802.22 WG on WRANs (Wireless Regional Area Networks) . IEEE. Retrieved January 18, 2009.

[4] ITU report ITU-R SM.2152 (09/2009), Definitions of Software Defined Radio (SDR) and Cognitive Radio System (CRS).

[5] Y. Saito et al., "Non-orthogonal multiple access (NOMA) for cellular future radio access", in Proc. IEEE 77thVeh.Technol. Conf., 2013, pp.1–5.

[6] D. Kim, H. Lee, and D. Hong, "A Survey of In-Band Full-Duplex Transmission: From the Perspective of PHY and MAC Layers", IEEE Comm. Surveys & Tutorials, vol. 17, no. 4, 4th quarter 2015, pp. 2017-2046.

[7] J. F. O'Hara and G. M. Moore, "A high performance CW receiver using feedthru nulling," Microw. J., vol. 6, no. 9, pp. 63–71, Sep. 1963.

[8] R. Ratasuk, D. Tolli , and A. Ghosh, "Carrier Aggregation in LTE-Advanced", In Proc. IEEE 71th Vehicular Technology Conference (VTC 2010-Spring), June 2011, pp 89 – 95.

[9] Y. Liao , L. Song, Z. Han, and Y. Li , "Full duplex cognitive radio: a new design paradigm for enhancing spectrum usage", IEEE Communications Magazine, Vol. 53, Issue 5, May 2015, pp 138 – 145.

[10] E. Askari , S. Aissa, "Full-duplex cognitive radio with packet fragmentation" In Proc. IEEE Wireless Communication and Networking Conference (WCNC) 2014.

[11] ET FCC. Docket no 03-222 notice of proposed rule making and order, Dec 2003.

[12] J. Mitola and G. Maguire, “Cognitive radio: making software radios more personal,” IEEE Personal Communications, vol. 6, no. 4, pp. 13– 18, Aug. 1999.

[13] A. Goldsmith and I. Marić, “Principals of Cognitive Radio Networks”, Chapter 2, Cambridge Univ. Press, 2012.

[14] A. Tonmukayakul and M. B. H. Weiss, “Secondary use of radio spectrum: A feasibility analysis,” in Proc. Telecommunications Policy Research Conference, Arlington, VA, USA, Oct. 2005.

[15] S. Shankar, C. Cordeiro, and K. Challapali, “Spectrum agile radios: utilization and sensing architectures,” in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Baltimore, Maryland, USA, Nov. 2005, pp. 160–169.

[16] Y. Yuan, P. Bahl, R. Chandra, P. A. Chou, J. I. Ferrell, T. Moscibroda, S. Narlanka, and Y. Wu, “KNOWS: Cognitive radio networks over white spaces,”, in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, Ireland, Apr. 2007, pp. 416–427.

[17] H. Tang, “Some physical layer issues of wide-band cognitive radio systems,” in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Baltimore, Maryland, USA, Nov. 2005, pp. 151–159.

[18] D. Cabric, S. Mishra, and R. Brodersen, “Implementation issues in spectrum sensing for cognitive radios,” in Proc. Asilomar Conf. on Signals, Systems and Computers, vol. 1, Pacific Grove, California, USA, Nov. 2004, pp. 772–776.

[19] T. Y¨ucek and H. Arslan, “Spectrum characterization for opportunistic cognitive radio systems,” in Proc. IEEE Military Commun. Conf., Washington, D.C., USA, Oct. 2006, pp. 1–6.

[20] J. Lehtom¨aki, M. Juntti, H. Saarnisaari, and S. Koivu, “Threshold setting strategies for a quantized total power radiometer,” IEEE Signal Processing Lett., vol. 12, no. 11, pp. 796–799, Nov. 2005.

[21] U. Gardner, WA, “Exploitation of spectral redundancy in cyclostationary signals,” IEEE Signal Processing Mag., vol. 8, no. 2, pp. 14–36, 1991.

[22] K. Maeda, A. Benjebbour, T. Asai, T. Furuno, and T. Ohya, “Recognition among OFDM-based systems utilizing cyclostationarity-inducing transmission,” in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, Ireland, Apr. 2007, pp. 516–523.

[23] P. D. Sutton, J. Lotze, K. E. Nolan, and L. E. Doyle, “Cyclostationary signature detection in multipath Rayleigh fading environments,” in Proc. IEEE Int. Conf. Cognitive Radio Oriented Wireless Networks and Commun. (Crowncom), Orlando, Florida, USA, Aug. 2007.

[24] D. Cabric and R. W. Brodersen, “Physical layer design issues unique to cognitive radio systems,” in Proc. IEEE Int. Symposium on Personal, Indoor and Mobile Radio Commun., vol. 2, Berlin, Germany, Sept. 2005, pp. 759–763.

[25] J. Lund´en, V. Koivunen, A. Huttunen, and H. V. Poor, “Spectrum sensing in cognitive radios based on multiple cyclic frequencies,” in Proc. IEEE Int. Conf. Cognitive Radio Oriented Wireless Networks and Commun. (Crowncom), Orlando, Florida, USA, July/Aug. 2007.

[26] M. Oner and F. Jondral, “Cyclostationarity based air interface recognition for software radio systems,” in Proc. IEEE Radio and Wireless Conf., Atlanta, Georgia, USA, Sept. 2004, pp. 263–266.

[27] A. Fehske, J. Gaeddert, and J. Reed, “A new approach to signal classification using spectral correlation and neural networks,” in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Baltimore, Maryland, USA, Nov. 2005, pp. 144–150.

[28] S. t. B. S. M. Mishra, R. Mahadevappa, and R. W. Brodersen, “Cognitive technology for ultra-wideband/WiMax coexistence,” in Proc. IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, Ireland, Apr. 2007, pp. 179–186.

[29] R. Tandra and A. Sahai, “Fundamental limits on detection in low SNR under noise uncertainty,” in Proc. IEEE Int. Conf. Wireless Networks, Commun. and Mobile Computing, vol. 1, Maui, HI, June 2005, pp. 464–469.

[30] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” IEEE J. Select. Areas Commun., vol. 3, no. 2, pp. 201–220, Feb. 2005.

[31] Z. Tian and G. B. Giannakis, “A wavelet approach to wideband spectrum sensing for cognitive radios,” in Proc. IEEE Int. Conf. Cognitive Radio Oriented Wireless Networks and Commun. (Crowncom), Mykonos Island, Greece, June 2006.

[32] T. Y¨ucek and H. Arslan, “A Survey of Spectrum Sensing Algorithms for Cognitive Radio Applications”, IEEE Communi. surveys & tutorials, vol. 11, no. 1, 1st quarter 2009. Pp. 116-130.

[33] E. Z. Tragos, S. Zeadally, A. G. Fragkiadakis, and V. A. Siris, “Spectrum Assignment in Cognitive Radio Networks: A Comprehensive Survey,” IEEE Commun. Surveys and tutorials, vol. 15, issue 3, third quarter 2013.pp 1108-1135.

[34] T.J. Harrold, L.F. Wang, M.A. Beach, G. Salami, A. Yarmohammad, and O. Holland, “Spectrum Sharing and Cognitive Radio: Opportunities for Efficiency Enhancement,” in Proc. International Workshop on Cognitive Wireless Communications and Networking 2009. St. Petersburg, Russia, 2009.

[35] D. Niyato, and E. Hossain,” Competitive spectrum sharing in cognitive radio networks: a dynamic game approach”, IEEE Trans. on wireless commun. Vol. 7, no. 7, July 2008.

[36] I.F. Akyildiz, W. Y. Lee, M.C. Vuran, and S. Mohanty,"A survey on spectrum management in cognitive radio networks," IEEE Communications Magazine, vol. 46, no. 4, pp. 40-48, Apr. 2008.

[37] P. Rawat, K.D.Singh, and J.M.Bonnin,“Cognitive radio for M2M and Internet of Things: A survey,” Elsevier Computer Communications journal, vol. 94, 15 Nov. 2016, pp. 1-29.

[38] Y. Zeng, N. Xiong, L. T. Yang, Y. Zhang, “Cross-layer Routing in Wireless Sensor Networks for Machine-to-Machine Intelligent Hazard Monitoring Applications,” IEEE Infocom conf. workshop on M2MCN-2011.

[39] C. Han, J. M. Jornet, E. Fadel, and I. F. Akyildiz, “A cross-layer communication module for the Internet of Things,” Elsevier Computer Networks Journal. vol. 57 (2013) pp. 622–633.

[40] X. Hong, J. Wang, C. Wang, and J. Shi, “Cognitive Radio in 5G: A Perspective on Energy-Spectral Efficiency Trade-off”, IEEE Communications Magazine, July 2014. Pp. 46-53.

[41] A. Raza, S. S. Ahmed, W. Ejaz, and H. S. Kim,” Cooperative Spectrum Sensing among Mobile Nodes in Cognitive Radio Distributed Network,” in proc. 10th International Conference on Frontiers of Information Technology, 2012.

[42] Y. Lin, H. Tembine, and K. Chen,” Inter-Operator Spectrum Sharing in Future Cellular Systems, ”, in proc. IEEE Globcom conf. 2012.

[43] H. Ju, D. Kim, H. V. Poor, and D. Hong, “Bi-directional beamforming and its capacity scaling in pairwise two-way communications,” IEEE Trans. Wireless Commun., vol. 11, no. 1, pp. 346–357, Jan. 2012.

[44] D. Kim, S. Park, H. Ju, and D. Hong, “Transmission capacity of full duplex based two-way ad-hoc networks with ARQ protocol,” IEEE Trans. Veh. Technol., vol. 63, no. 7, pp. 3167–3183, Sep. 2014.

[45] H. Ju, S. Lim, D. Kim, H. V. Poor, and D. Hong, “Full duplexity in beamforming-based multi-hop relay networks,” IEEE J. Sel. Areas Commun., vol. 30, no. 8, pp. 1554–1565, Sep. 2012.

[46] M. Jain et al., “Practical, real-time, full duplex wireless”, in proc. the 17th annual international conference on Mobile computing and networking (Mobicom) 2011, September, pp. 301-312.

[47] G. Zheng, I. Krikidis, J. Li, A. P. Petropulu, and B. Ottersten, “Improving physical layer secrecy using full-duplex jamming receivers,” IEEE Trans. Signal Process., vol. 61, no. 20, pp. 4962–4974, Oct. 2013.

[48] M. Duarte et al., “Design and characterization of a full-duplex multi antenna system for WiFi networks,” IEEE Trans. Veh. Technol., vol. 63, no. 3, pp. 1160–1177, Mar. 2014.

[49] A. Sabharwal et al., “In-band full-duplex wireless: Challenges and opportunities”, IEEE J. Sel. Areas Commun., vol.32,no.9,pp.1637–1652, Sep. 2014.

[50] D. Bharadia, E. McMilin, and S. Katti, “Full duplex radios,” in Proc. ACM SIGCOMM, 2013, pp. 375–386.

[51] M. Duarte and A. Sabharwal, “Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results,” in Proc. 44th Asilomar Conf. Signals, Syst. Comput., 2010, pp. 1558–1562.

[52] T. Riihonen, S. Werner, and R. Wichman, “Spatial loop interference suppression in full-duplex MIMO relays,” in Proc. 43rd Asilomar Conf. Signals, Syst. Comput., 2009, pp. 1508–1512.

[53] IEEE Journal on Selected Areas in Communications, Special Issue on “Advances in Cognitive Radio Networking and Communications - I”, vol. 29, no. 2, February 2011.

[54] IEEE Journal on Selected Areas in Communications, Special Issue on “Advances in Cognitive Radio Networking and Communications - II”, vol. 29, no. 4, April 2011.

[55] W. Cheng, Xi Zhang, and H. Zhang, “Full duplex wireless communications for cognitive radio networks”, arXiv:1105.0034 [cs.IT], April 2011.

[56] Y. Liao, L. Song, Z. Han, and Y. Li, “Full-Duplex cognitive radio: a new design paradigm for enhancing spectrum usage”, arXiv:1503.03954 [cs.NI], March 2015.

[57] L. T. Tan and L. B. Le, “Distributed MAC protocol design for full–duplex cognitive radio networks”, arXiv:1506.08328 [cs.IT], June 2015.

[58] W. Cheng, X. Zhang, and H. Zhang, “Full duplex spectrum sensing and MAC protocol for multichannel non time slotted cognitive radio networks”, IEEE Journal on Sel. Areas in Communications.

[59] G. Zheng, I. Krikidis, and B. Ottersten,” Full duplex cooperative cognitive radio with transmit imperfections”, IEEE Trans. on Wireless Comm. vol. 12, no. 5, May 2013.

[60] E. Ahmed, A. Eltawil, and A. Sabharwal, “Simultaneous transmit and sense for cognitive radios using full-duplex: A first study”, in Proc. IEEE Antennas Propag. Soc. Int. Symp., 2012, pp. 1–2.

[61] E. Tsakalaki, O. N. Alrabadi, A. Tatomirescu, E. de Carvalho, and G. F. Pedersen, “Concurrent communication and sensing in cognitive radio devices: Challenges

and an enabling solution," IEEE Trans. Antennas Propag., vol. 62, no. 3, pp. 1125–1137, Mar. 2014.

[62] J. Heo, H. Ju, S. Park, E. Kim, and D. Hong, "Simultaneous sensing and transmission in cognitive radio," IEEE Trans. Wireless Commun., vol. 13, no. 4, pp. 1948–1959, Apr. 2014.

[63] W. Afifi and M. Krunz, "Incorporating self-interference suppression for full-duplex operation in opportunistic spectrum access systems," IEEE Trans. Wireless Commun., vol. 14, no. 4, pp. 2180–2191, Apr. 2015.

[64] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in Proc. IEEE INFOCOM, 2013, pp. 1258–1266.

[65] W. Afifi, and M. Krunz, "Adaptive transmission-reception-sensing strategy for cognitive radios with full-duplex capabilities", in Proc. of the IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN), 2014, April 2014, pp 149-160.

[66] O. Simeone et al., "Spectrum leasing to cooperating secondary ad-hoc networks," IEEE J. Sel. Areas Commun., vol. 26, no. 1, pp. 203–213, Jan. 2008.

[67] Y. Liao, T. Wang, L. Song, and Z. Han," Listen-and-Talk: Protocol Design and Analysis for Full-duplex Cognitive Radio Networks", IEEE Trans. on Vehicular Technology, Feb 2016.

[68] G. Zheng, I. Krikidis, and B. Ottersten," Full-Duplex Cooperative Cognitive Radio with Transmit Imperfections", IEEE Trans. on wireless Commun., vol. 12, no. 5, May 2013.

[69] G. Zheng, E. A. Jorswieck, and B. Ottersten, "Cooperative communications against jamming with half-duplex and full-duplex relaying," in Proc. IEEE 77th Veh. Technol. Conf., 2013, pp. 1–5.

[70] J. Palicot, J. Mitola III, Z. Lei, and F.K. Jondral, "10 years of cognitive radio: state-of-the-art and perspectives", EURASIP Journal on Wireless Communications and Networking, Special Issue, July 2012.

[71] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty, “Next generation dynamic spectrum access cognitive radio wireless networks: a survey,” Computer Networks, vol. 50, no. 13, pp. 2127–2159, 2006.

[72] A.R. Haqiqatnejad, K. Shahtalebi, and A.R. Forouzan,” Spectrum Leasing with Disjoint Secondary User Selection in Cognitive Radio Networks”, in proc. IEEE Workshop on Communication and Information Theory (IWCIT), Iran, 2014.

[73] J. C. F. Li, W. Zhang, A. Nosratinia, and J. Yuan, “SHARP: Spectrum harvesting with ARQ retransmission and probing in cognitive radio”, IEEE Trans. on Comm., vol. 61, no. 3, Mar 201, pp. 951-960.

[74] K. Eswaran, M. Gastpar, and K. Ramchandran, “Bits through ARQs: spectrum sharing with a primary packet system,” in Proc. 2007 IEEE Int. Symp. on Information Theory, pp. 2171–2175.

[75] ——, “Cognitive radio through primary control feedback,” IEEE J. Sel. Areas Commun., vol. 29, no. 2, pp. 384–393, Feb. 2011.

[76] M. Levorato, U. Mitra, M. Zorzi, “Cognitive Interference Management in Retransmission-Based Wireless Networks”, IEEE Transactions on Information Theory, Vol 58, pp 3023 – 3046, May 2012.

[77] R. A. Tannious and A. Nosratinia, “Cognitive radio protocols based on exploiting hybrid ARQ retransmissions,” IEEE Trans. Wireless Comm. vol. 9, no. 9, pp. 2833–2841, Sep. 2010.

[78] Peterson and Davie, “Computer Networks: A Systems Approach”, Third Edition, 2003.

[79] N. Michelusi, O. Simeone, M. Levorato, P. Popovski, and M. Zorzi, “Optimal cognitive transmission exploiting redundancy in the primary ARQ process,” in Proc. 2011 Information Theory and Applications Workshop, pp. 1–10.

[80] N. Michelusi, P. Popovski, M. Levorato, O. Simeone, and M. Zorzi, “Cognitive transmissions under a primary ARQ process via backward interference cancellation,” in Proc. 2011 Allerton Conf. on Commun. Control and Computing, pp. 727–735.

[81] M. Levorato, U. Mitra, and M. Zorzi, “Cognitive Interference Management in Retransmission-Based Wireless Networks”, IEEE Trans. on Information Theory, vol. 58, pp 3023 – 3046, May 2012.

[82] X. Wang, Y. Ji and J. Li, “Cooperative coding based retransmission protocol for cognitive radio networks by exploiting hybrid ARQ”. In Proc. of Wireless Communications and Mobile Computing Conference (IWCMC) 2014, pp 399-404.

[83] S. Zhang, S. C. Liew, L. Lu, and H. Wang, “Blind known interference cancellation with parallel real valued belief propagation algorithm”, in Proceedings of IEEE Globecom Conf. 2012, Dec. 2012.

[84] S. M. Alamouti,” A Simple Transmit Diversity Technique for Wireless Communications”, IEEE J. on Selected areas in communications, vol. 16, no. 8, Oct 1998, pp. 1451-1458.

[85] J. G. Proakis, Digital communications, 5th Edition, McGraw-Hill, 2007.

[86] "IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE-SA. 5 April 2012.

[87] Q. Liu, S. Zhou, and G. B. Giannakis, “Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links”, IEEE Trans. on Wireless Comm., vol. 3, no. 5, Sept. 2004, pp. 1746-1755.

[88] S. Huang, X. Liu, and Z. Ding, “Optimal sensing-transmission structure for dynamic spectrum access”, in Proc. of the IEEE Infocom’09 Conf., April 2009, pp. 2295–2303.

[89] W. Afifi, A. Sultan, and M. Nafie, “Adaptive sensing and transmission durations for cognitive radios,” in Proc. of the IEEE DySPAN’11 Conf., 2011, pp. 380–388.

[90] I.F. Akyildiz, B.F. Lo, and R. Balakrishnan, “Cooperative spectrum sensing in cognitive radio networks: A survey”, Journal Physical Comm., vol. 4, no. 1, March 2011, pp 40-62.

[91] W. Zhang, R. K. Mallik, and K. B. Letaief, “Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks”, IEEE Trans. on Wireless Commun., vol. 8, no. 12, pp 5761–5766, Dec 2009.

[92] J. Unnikrishnan, and V. V. Veeravalli, “Cooperative sensing for primary detection in cognitive radio”, IEEE Journal of selected topics in signal processing, vol. 2, no. 1, Feb 2008.

[93] A. Sahai, G. Patel, and A. Sabharwal, “Asynchronous full-duplex wireless”, in Proc. Of the 4th International Conference on Communication Systems and Networks (COMSNETS 2012), Jan. 2012, pp 1-9.

[94] Y.C. Liang, Y. Zeng, E.C.Y. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks”, IEEE Trans. on Wireless Comm., vol. 7, no. 4, Apr 2008.

[95] Y. Song, S. D. Blostein, and J. Cheng,” Exact outage probability for equal gain combining with co-channel interference in Rayleigh fading”, IEEE Trans. On Wireless Commun., vol. 2, no. 5, Sept. 2003, pp. 865-870.

[96] F. Ramos-Alarcon, V. Kontorovich, and M. Lara,” Outage probability in nakagami channels using fade duration distribution approximate results,” IEEE Trans. on Commun., vol. 58, no. 4, Apr. 2010, pp. 1009-1013.

[97] J. Lai, and N. B. Mandayam, “Minimum duration outages in Rayleigh fading channels”, IEEE Trans. on Commun., vol. 49, no. 10, Oct 2001, pp. 1755-1761.

[98] S. O. Rice, “Distribution of the duration of fades in radio transmissions: Gaussian noise model,” Bell Syst. Technol. J., vol. 37, pp. 581–635, May 1958.

[99] Y. Chen, Q. Zhao, and A. Swami, “Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors,” IEEE Trans. on Info. Theory, vol. 54, no. 5, pp. 2053–2071, May 2007.

[100] Y.C. Liang, Y. Zeng, E.C.Y. Peh, and A. T. Hoang, “Sensing-throughput tradeoff for cognitive radio networks”, IEEE Trans. on Wireless Comm., vol. 7, no. 4, Apr 2008, pp. 1326-1337.

[101] W. Y. Lee and I. F. Akyildiz, “Optimal spectrum sensing framework for cognitive radio networks,” IEEE Trans. on Wireless Comm., vol. 7, no. 10, pp. 3845–3857, Oct. 2008.

[102] Y. Chen, Q. Zhao, and A. Swami, “Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors,” IEEE Trans. on Info. Theory, vol. 54, no. 5, pp. 2053–2071, May 2007.

[103] D. W. Bliss, P. Parker, and A. R. Margetts, “Simultaneous transmission and reception for improved wireless network performance,” in Statistical Signal Processing, 2007. SSP ’07. IEEE/SP 14th Workshop on, Aug. 2007, pp. 478–482.

[104] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, ”Achieving single channel full duplex wireless communications”, in Proc. ACM MobiCom, 2010, pp. 1–12.

[105] P. Larsson and M. Prytz, “MIMO on-frequency repeater with self-interference cancellation and mitigation,” in Proc. IEEE 69th VTC—Spring, 2009, pp. 1–5.

[106] E. Aryafar, M. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, “MIDU: Enabling MIMO full duplex,” in Proc. ACM MobiCom, 2012, pp. 257–268.

[107] S. Zhang, S. C. Liew, and P. P. Lam, “Physical-layer network coding,” in Proc. ACM MobiCom, 2006, pp. 358–365.

[108] Z. Li, M. Peng, and W. Wang, “A network coding scheme for the multiple access full-duplex relay networks,” in Proc. 6th Int. ICST Conf. Commun. Netw. China, 2011, pp. 1132–1136.

[109] E. Ahmed, A. M. Eltawil, and A. Sabharwal, “Self-interference cancellation with nonlinear distortion suppression for full-duplex systems,” CoRR, vol. abs/1307.3796, 2013.

[110] E. Ahmed, A. Eltawil, and A. Sabharwal, “Rate gain region and design tradeoffs for full-duplex wireless communications,” IEEE Trans. Wireless Commun., vol. 12, no. 7, pp. 3556–3565, Jul. 2013.

[111] O. Somekh, O. Simeone, H. Poor, and S. Shamai, “Cellular systems with full-duplex amplify-and-forward relaying and cooperative basestations,” in Proc. IEEE Int. Symp. Inf. Theory, 2007, pp. 16–20.

[112] B. Day, A. Margetts, D. Bliss, and P. Schniter, “Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range,” IEEE Trans. Signal Process., vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[113] Z. Li, X.-G. Xia, and M. Lee, “A simple orthogonal space-time coding scheme for asynchronous cooperative systems for frequency selective fading channels,” IEEE Trans. Commun., vol. 58, no. 8, pp. 2219–2224, Aug. 2010.

[114] M. Samavat, F. Hosseini, and S. Talebi, “Alamouti coding scheme for cooperative relay networks with full duplex relaying,” in Proc. Iran Workshop Commun. Inf. Theory, 2013, pp. 1–4.

[115] M. J. Emadi, A. G. Davoodi, and M. R. Aref, “Analytical power allocation for a full duplex decode-and-forward relay channel,” IET Commun., vol. 7, no. 13, pp. 1338–1347, Sep. 2013.

[116] V. Towhidlou, M. Shikh-Bahaei, O. Holland, M. Dohler, and H. Aghvami, ”Directional antennas; full duplex communication” contribution submitted to IEEE 802.22 standard committee, doc.: IEEE 802.22-16/0033r0, Nov. 2016.

[117] J. Choi, S. Kam, D. Kim, and D. Hong, “Outage probability of full duplex systems in multi-spectrum environment,” in Proc. 4th Int. Conf. Adv. Commun. Comput., 2014, pp. 65–69.

[118] R. Gallager, “Information Theory and Reliable Communication,” New York: Wiley, 1968.

[119] D. Wu and R. Negi, “Effective capacity: a wireless link model for support of quality of service,” IEEE Trans. Wireless Comm., vol. 2, no. 4, pp. 630-643, July 2003.

[120] L. Kleinrock, “Queuing Systems, Volume I: Theory” New York, Wiley, 1975.

[121] A Shojaeifard, F Zarringhalam, and M Shikh-Bahaei, “Joint physical layer and data link layer optimization of CDMA-based networks” IEEE Trans. on Wireless Communications, vol. 10, no. 10, pp. 3278-3287, Aug 2011.

[122] H. Bobarshad, M. van der Schaar, and M. R. Shikh-Bahaei, “A low complexity analytical modelling for cross-layer adaptive error protection in video over WLAN,” IEEE Trans. Multimedia, vol. 12, no. 5, pp.401–414, Apr. 2012.

[123] H. She, Z. Lu, A. Jantsch, D. Zhou, and L. Zheng, “Modeling and Analysis of Rayleigh Fading Channels using Stochastic Network Calculus” in proc. IEEE WCNC conf. 2011.

[124] S. Hanly and D. Tse, “Multiaccess Fading Channels - Part II: Delay-Limited Capacities,” IEEE Trans. Info. Theory, v. 44, no. 7, Nov. 1998, pp. 204-221.

[125] E. Teletar and R. G. Gallager, “Combining Queueing Theory with Information Theory for Multiaccess” IEEE Jour. Sel. Areas in Comm., vol.13, no.6, Aug. 1995, pp. 963-969.

[126] W. Turin and M. Zorzi, “Performance Analysis of Delay-Constrained Communications Over Slow Rayleigh Fading Channels” IEEE Trans. on Wireless Comm., vol. 1, no 4, Oct 2002, pp. 801-807.

[127] J. A. Ritcey, “Prony Approximations for Probability density functions”, in proc. of IEEE Pacific Rim Conference on Communications, Computers, and Signal Processing, pp. 185-188, 1995.

[128] R. F. Botta, “Approximation with Generalized Hyperexponential Distributions: Weak Convergence Resulats,” Springer Queuing Systems, vol 1, Issue 2, pp 169-190, Sept 1986.

[129] Y. Dong and P. Fan, “Queueing Analysis for Block Fading Rayleigh Channels in the Low SNR Regime,” in proc. IEEE International Conference on Wireless Communications & Signal Processing (WCSP) 2013, pp. 1-6, Oct. 2013.

[130] L. Chang, Q. Xiaowei, Z. Sihai, Z. Wuyang,” Queuing Analysis of Self-similar Traffic in Rayleigh Fading Channel,” in proc. International Conference on Networks Security, Wireless Communications and Trusted Computing, 2009.

[131] G. Verticale, “A Closed-Form Expression for Queuing Delay in Rayleigh Fading Channels Using Stochastic Network Calculus” in proc. of the 5th ACM symposium on QoS and security for wireless and mobile networks, pp 8-12, 2009.

[132] Y. Dong, Q. Wang, P. Fan and K. B. Letaie, “The Deterministic Time-Linearity of Service Provided by Rayleigh Fading Channels,“ in proc. Wireless Comm. and Mobile Computing Conference (IWCMC), pp. 1519-1523, July 2011.